Vrije Universiteit Brussel

FACULTY OF ENGINEERING
Department of Fundamental Electricity and Instrumentation

Model Fitting in Frequency Domain Imposing Stability of the Model

Thesis submitted in fulfillment of the requirements for the degree of Doctor in de ingenieurswetenschappen (Doctor in Engineering) by

ir. László Balogh

Advisers: Prof. dr. ir. Rik Pintelon
            Prof. dr. ir. István Kollár

October 2009

Vrije Universiteit Brussel

FACULTY OF ENGINEERING
Department of Fundamental Electricity and Instrumentation

Model Fitting in Frequency Domain Imposing Stability of the Model

Thesis submitted in fulfillment of the requirements for the degree of Doctor in de ingenieurswetenschappen (Doctor in Engineering) by

ir. László Balogh

Chairman: Prof. dr. ir. Patrick Guillaume

Vice chairman: Prof. dr. Steve Vanlanduit

Secretary: Prof. dr. ir. Johan Schoukens

Promoters: Prof. dr. ir. Rik Pintelon
                   Prof. dr. ir. István Kollár

Jury: Prof. dr. ir. Joos Vandewalle
                   Prof. dr. Barnabás Garay

October 2009

Contents

Abstract

László Balogh
"Model Fitting in Frequency Domain Imposing Stability of the Model"
PhD thesis

Most signal processing methods can be formulated equivalently in both time and frequency domains. In this thesis we focus only on frequency domain algorithms. Parts of two signal processing topics are investigated: approximation and identification problems. Although they use very different tools and have different goals, still, there are similarities.

The two problems are in fact consecutive steps of an identification procedure. The first step is the unconstrained estimation of the model parameters from noisy measurement data. The result is a validated model with uncertainty bounds. If the validated model is unstable, then the identification step is followed by an approximation step where the validated model is approximated by a stable one. The final result is a stable model with bias error and uncertainty bounds. The topics of this thesis are the stable approximation and the identification method of linear dynamic systems.

The contribution to the first step is the following. To improve the numerical stability of identification algorithms the parameter space is usually transformed. If any constraint of the parameter space is imposed then it must be transformed, too. In the case of so-called total least squares estimator a theoretical gap is filled by showing how the constraints can be transformed. Variations of the method are also analyzed.

Next, contributions are made to the second step of the identification procedure. In simulation and prediction applications unstable models are usually undesirable. To overcome this problem the unstable transfer function must be approximated by a stable one. Approximation theory has rich literature, nevertheless there are lots of open questions. A natural solution to this problem is imposing constraints on the poles such that they are guaranteed to lie in the stability region. Another possible solution consists in adding an appropriate delay to the target function. Its main disadvantage is that it can be used for open-loop configurations only. But a strong benefit is that the resulting parameter vector of the minimization algorithm determines a stable model without imposing any constraint.

A practical and a theoretical contribution in the field of the stable approximation are presented in the thesis. This dissertation presents a theorem which guarantees the existence of a delay added to a target function for second order systems such that the approximation is stable. Under some reasonable assumptions the theorem is extended to systems of arbitrary order. Moreover, a new numerical algorithm which is capable to find the appropriate stable approximation is proposed and analyzed. The thesis also compares the results with those found in the literature.

Kivonat (Abstract in Hungarian)

Balogh László
„Nem stabil átviteli függvények stabil approximációja"
PhD értekezés

A jelfeldolgozási módszerek idő vagy frekvencia tartományban működnek. A legtöbb módszer mindkét tartományban felírható ekvivalens módon, esetünkben frekvencia tartománybeli algoritmusokat tárgyalunk. A dolgozatban két témakörről találhatóak eredmények: stabil approximációs problémákról és egy identifikációs módszerről.

A két probléma az identifikációs folyamat egymás követő lépései. Először zajos mérési adatokból a modell paraméter becslését végezzük el. Az eljárás eredménye egy validált modell és a modell paramétereinek szórása. Lehetséges, hogy a modell nem stabil, ezért egy stabil approximációs eljárást alkalmazunk. A végső eredménye egy stabil modell, valamint a modell paramétereinek szórás, illetve a torzításának felső becslése.

Az identifikációs folyamat első lépéshez a dolgozat következő eredménye tartozik. Az identifikációs módszerek alkalmazása során sokszor transzformáljuk a paraméter vektorok terét, hogy a számítási pontosságot és a numerikus stabilitást növeljük. Amennyiben a paraméter vektorra feltételekkel is megfogalmazunk, akkor azokat is transzformálni kell. Az úgy nevezett total least squares becslő esetében megmutatjuk, hogyan kell a feltételeket transzformálni. A dolgozat a módszer további változatait is elemzi.

Az identifikáció folyamat második lépéséhez tartozó eredmények a stabil approximáció témakörébe tartoznak. A gyakorlati alkalmazásokban a nem stabil modellek általában nem használhatóak. A probléma egyik megoldása, ha a nem stabil átviteli függvényt egy stabillal approximáljuk. Az approximációs elméletnek gazdag irodalma van, mindemellett sok nyitott kérdés van. Egy lehetséges megoldása a stabil approximációnak lineáris, időinvariáns rendszerek esetén az, ha a célfüggvényhez késleltetést adunk. Ennek a legnagyobb hátránya az, hogy csak nem visszacsatolt rendszerek esetén tudjuk az eredményt használni. Ám egy nagy előnye a megközelítésnek, hogy a költségfüggvény minimuma egy stabil rendszert határozhat meg anélkül, hogy bármilyen megkötést alkalmaznánk a minimalizáló eljárás során.

A disszertációban bebizonyítunk egy tételt, amely kimondja, hogy ha elég nagy késleltetést adunk bármilyen másodfokú átviteli függvényhez, akkor az approximáció négyzetes költségfüggvényének minimuma egy stabil rendszert határoz meg. Az általános esetet bebizonyítjuk olan feltételek mellett, amelyeket gyakorlati tapasztalatok támasztanak alá. Tovább bemutatunk és megvizsgálunk egy olyan eljárást, amely képes megtalálni az említett stabil illesztést. A dolgozat összehasonlító elemzést is végez az irodalomban fellelhető módszerekkel.

Glossary of Symbols

Symbols

A	Data matrix of the TLS problem.
A_m	Corrected data matrix of the TLS problem.
$\angle$	Angle of a complex number.
$\mathbb{C}$	Complex plane.
C	Covariance of A.
C_m	Covariance of A_m.
$D(\Omega, \theta)$	Denominator of $H(\Omega, \theta)$.
$e^{j\omega}$	Complex frequency in z-domain.
$\mathbb{E}$	Expected value operator.
F	Number of frequencies.
$H_0(\Omega)$	Real transfer function in estimation problems.
$H(\Omega, \theta)$	Estimation or approximation of the transfer function.
Im{}	Imaginary part of a complex number.
L_2	Space of the square integrable functions.
μ	Lebesgue measure.
n_α	Order of the denominator.
n_β	Order of the numerator.
$N(\Omega, \theta)$	Numerator of the transfer function $H(\Omega, \theta)$.
$j\omega$	Complex frequency in s-domain.
Ω	Generalised frequency.
$\mathbb{R}$	Real line.
Re{}	Real part of a complex number.
$\mathcal{R}(A)$	Range of the matrix A.
Res	Residue of a complex function.
$R_r(x)$	Fortshyte orthogonal polynomials.
$\mathbb{T}$	Unit circle.
τ	Delay.
θ	Parameter vector.
$T(\Omega)$	Target function in approximation problems.
z	Complex variable.

Abbreviations

BL	Band Limited
BIBO	Bounded-Input Bounded-Output
dB	decibel
DFT	Discrete Fourier Transform
DUT	Device Under Test
FRF	Frequency Response Function
GSVD	Generalised Singular Value Decomposition
GTLS	Generalised TLS
Hz	Hertz
LS	Least Squares
ML	Maximum Likelihood
ODE	Ordinary Differential Equations
SVD	Singular Value Decomposition
TLS	Total Least Squares
ZOH	Zero-Order-Hold

Preface

This thesis is the collection of the main results achieved during my research conducted at Vrije Universiteit Brussel (VUB), Brussel, Belgium and at the Budapest University of Technology and Economics (BUTE), Budapest, Hungary.

I have started my PhD studies in 2000, right after graduation, at the Department of Measurement and Information Systems (BUTE) under the supervision of Prof. István Kollár. In period from 2001 till 2006 I spent altogether about 10 months as a visiting researcher at Vrije Universiteit Brussel, where I was working under the supervision of Prof. Rik Pintelon.

The first result was the correction of the total least squares estimator. After that, during my first visit in Brussels Rik Pintelon proposed me a related research question. So, I started to think about some problems of stable approximation with additional delay. Because of my other studies the research took several years, but in the end I could prove an existence theorem and a new algorithm was proposed.

Collecting the results I started to write the thesis in 2005 when I was in Brussels. After returning to Hungary in April 2006 I have been writing this work.

I would like to extend my thank to my thesis advisors, Prof. Rik Pintelon and Prof. István Kollár who have proposed the topics of my thesis work and helped me with their invaluable guidance and constant motivation.

I am most grateful for the many formal and informal professional discussion that I could have held with colleagues of both departments. I thank Johan Schoukens, Yves Rolain and Tom D'haene from Belgium and János Márkus, József Németh, Tamás Bakó, Balázs Vödrös, Balázs Vargha, László Sujbert, Zoltán Szabó, Károly Molnár and Attila Sárhegyi from Hungary for their help and the useful discussions. Also many thanks to Bertalan Eged, who constantly supported my work.

Special thanks to my best friend Tamás Priskin, who has been always ready to help me without questions. Last but not least I would also like to thank my family: Timi (my wife), György (my brother) and my parents for their constant love and support. Without them I would have never succeed.

Chapter 1

Introduction

1.1 A Note of Clarification

The thesis presents results from two separate fields in system identification, which are related by the fact that both were part of my doctoral work. They mirror the fact that my involvement had changed along the way. The order of the topics is reversed compared to the chronological order.

In this work a two-step identification procedure is presented. In the first step, an unconstrained model is identified from measured data. The result of this step is a validated model with uncertainty bounds. If the validated model is unstable, then an additional step is performed. In the second step the unstable validated model is approximated by a stable one and the result is a stable model with bias error bounds. The end result of both steps is a stable model with bias error and uncertainty bounds. In this thesis problems from both steps are investigated and solved. The correction of the total least square method belongs to problems from the first step of the identification procedure. The contribution to the second step is a special algorithm which solves the stable approximation problem by adding a delay to the target function. A theoretical result and a practical algorithm are presented.

1.2 Stable Approximation with Additional Delay

The increase of the computation speed and precision of computers allows to solve more complex problems than before. However, without proper algorithms the computational complexity cannot be decreased. For several digital signal processing problems it is easy to formulate the problem but it is very hard to solve it. The stable approximation problem in frequency domain is one of the examples.

Approximation and design are closely related. Designing a filter always means that a parameter vector which is the best in some sense should be determined. And it is exactly the same in approximation problems. Hence, we will use these terms interchangeably. Filter designs can be done in time or in frequency domain. Filter specification is more common in the frequency domain because design constraints can be directly formulated by using simple functions of the frequency domain. Stable approximation is a design

technique in which the ultimate goal of the approximation approach is to obtain a stable model. The theory and the practise of approximation is very rich and there is no chance to put everything in a thesis with limited size. The models or the systems in this work are linear and time invariant, [Oppenheim et al., 1983]. In spite of the numerous theorems about these systems there are many open questions.

Finite impulse response (FIR) filters are always stable. Stability problems may occur during the design of an infinite impulse response (IIR) filter. To overcome this problem several methods were proposed. The pole-constrained IIR filter design method is a modification of the well-known Ellacott-Williams algorithm [Ellacott and Williams, 1976b] which is based on [Deczky, 1972]. Unfortunately, in lots of cases imposing constraints leads to bad approximation. Hence, in some cases other methods must be used. Practical examples show that if a delay is added to the target function, then the final model which is the result of the approximation might be stable. An intuitive description of this phenomenon was presented in [Johnson, 1978]. However, a strict mathematical proof was not available.

[Vuerinckx, 1998] presents a new practical algorithm which can find a delay that must be added to the target function such that the obtained parameter vector determines a stable model. Although for all examples studied in [Vuerinckx, 1998] it was found that adding a delay to the target function stabilizes the approximation, no exact mathematical proof of this conjecture was given.

In this thesis a theorem is proved which guarantees for second order systems that if enough delay is added to the target function, then the approximation is stable. The proof is valid for a continuous frequency interval only, but there is a same consequence for a finite frequency grid because arbitrary good approximations of the corresponding integrals can be obtained by increasing the dense of the frequency grid. For higher order systems the proof is extended using reasonable assumptions which are based on practical experience with many approximation/filter design problems. Moreover, a new practical algorithm is presented to find the optimal delay value.

1.3 Correction of the Total Least Squares Method

The so-called least squares (LS) method is very popular and used in lots of applications. In the case of the LS method we assume that only the right hand side of the normal equations is noisy. A possible extension is the total least squares (TLS) method in which it is supposed that both the right hand side and the system matrices are corrupted by noise [Van Huffel and Vandewalle, 1991]. The TLS method has several applications, too. This thesis focuses on the TLS method in frequency domain system identification.

Identification is a statistical method to determine parametric or non-parametric variables of a system. Like everything from digital signal processing, there are methods which are working in the time-domain and algorithms which are processing frequency domain data. The time domain methods are summarized in [Ljung, 1999]. Methods of the frequency domain system identification are described firstly in a book in 1991, [Schoukens and Pintelon, 1991]. After more than 10 year long research a new book was published,

[Pintelon and Schoukens, 2001]. The second book already contains the introduction and the analysis of the TLS method in frequency domain system identification.

Through an identification approach, the space of the possible parameter values is usually transformed. The aim might be to increase the numerical stability. A subset of this kind of transformations is the linear mapping which is nothing else than a basis transformation. The TLS estimation is equivalent to solving of a set of linear equations and imposing a constraint on the length of the parameter vector. The second part of the thesis is about the transformation of the constraint.

1.4   Structure of the Thesis

Every chapter was written such that it can be read independently. Cross-relation between chapters was unavoidable but endeavored to be minimized.

Chapter 2 introduces the frequency domain. The Fourier transformation is defined for the L_1 and L_2 cases. Moreover, the so-called Hardy spaces are presented. After that, the subset of models which are used in the thesis can be found. This chapter is the common ground of the stable approximation and the identification in frequency domain.

Chapter 3 reviews the already existing and published method pertaining to stable approximation and identification problems. It contains not only stabilization procedures with additional delay but also methods in which no additional delay is allowed.

The complete chapter 4 is about an existence theorem and its consequences. The theorem says roughly that if in the case of low order systems enough delay is added to the target function, then the result of the optimization without imposing any constraint determines a stable model.

After that, in chapter 5 a numerical algorithm is presented. Numerical tools of ordinary differential equations are applied to develop a new search algorithm. Examples are shown to present its effectiveness.

The topic of the chapter 6 is the exploration of the TLS method which is a very effective tool in frequency domain system identification, and showing an extension which enables us to discuss the transformations in a common framework.

Finally, chapter 7 presents conclusions and suggests possible directions for future research.

Chapter 2

The Frequency Domain

The object of this chapter is to introduce the basic tools used in this thesis in which an approximation problem and an identification problem are investigated deeply in the frequency domain. In contrast to the time domain in this world representations of real quantities, function, etc. are complex. The algorithms and the problems differ from the ones in the time domain. Therefore the main aim of this chapter is to clarify the mathematical background and to introduce some notations which are in the intersection of studied identification and approximation problems.

First, the Fourier transformation is introduced. The importance of this tool is well known. Here different versions are studied and the most important properties are described. In the next section the so-called Hardy spaces are investigated and the origin and the basics of approximation problems in the frequency domain are established. The third section is about the models used throughout identification and approximation in the frequency domain. After that, a short description about the similarities of and the differences between identification and approximation is given.

2.1 Fourier Transform

The Fourier transform is a common concept in mathematics. This is a tool of complex analysis, functional analysis, differential equations and so on.

Definition 2.1.1 (Fourier Transform, L_1 case)**.** *Let $f(x) : \mathbb{R} \mapsto \mathbb{C}$ denote a function from the space L_1 then the Fourier transform $\mathcal{F}(f) = F(y)$ of this function is*

$$\mathcal{F}(f) = F(y) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi xy}dx, \tag{2.1}$$

where $y \in \mathbb{R}$, $j = \sqrt{-1}$ and $F(y) \subset \mathbb{C}$.

Informally, a L_1-function is a function $f : \mathbb{R} \mapsto \mathbb{C}$ that is absolute integrable, i.e.

$$\int_{-\infty}^{\infty} |f(x)|dx < \infty, \tag{2.2}$$

where the integral is in Lebesgue sense. It can be seen that the function $f(x)$ and the image function $F(y)$ have the same domain ($\mathbb{R}$) and the same image ($\mathbb{C}$).

However, sometimes functions are in the L_2 space. The $L_2(I)$ with $I \subset \mathbb{R}$ denote the space of all square integrable and Lebesgue measurable functions

$$L_2(I) = \left\{ f(x) : \int_I |f(x)|^2 dx < \infty \right\} \tag{2.3}$$

and the norm on this space is

$$\| f \|_2 = \sqrt{\int_I f(x)f^*(x)dx}, \tag{2.4}$$

where star denotes the complex conjugation. This norm can be inherited from a scalar product which is by definition

$$< f, g > := \int_I f(x)g^*(x)dx \tag{2.5}$$

for $f, g \in L_2(I)$. The definition of L_2 is $L_2 = L_2(-\infty, \infty)$.

The Fourier transformation can be extended to this case [Titchmarsh, 1937]. Without going into details this extension is by definition the same with (2.1) if $f(x) \in L_1 \cap L_2$. One of the main advantages of the function space L_2 is that it is a so called Hilbert space [Rudin, 1991], so the norm can be inherited from a scalar product. In the background in every engineering application the Fourier transform of the space L_2 is used because of its good properties. For example, the Parseval relation can be defined in the space L_2 only. Moreover, the image functions of the space L_2 constitutes an L_2 space again. In the case of the L_1 space the Fourier transform of a uniform impulse function can be defined but the image which is a sinc function is not in L_1. This is not a problem in the case of L_2 space.

Summarizing the properties of the L_2 Fourier transform:

- Its domain has more importance in the practical functions.

- It is a linear, unitary operation $\mathcal{F} : L_2 \mapsto L_2$. (A unitary operator is a bounded linear operator $\mathcal{F}$ on a Hilbert space satisfying $\mathcal{F}\mathcal{F}^* = \mathcal{F}^*\mathcal{F} = $ Identity where $\mathcal{F}^*$ is the adjoint of $\mathcal{F}$.)

- The scalar product operation is preserved (Parseval relation):

$$< \mathcal{F}(f), \mathcal{F}(g) > = < f, g > . \tag{2.6}$$

Definition 2.1.2 (Discrete Time Fourier transformation). *If the measure in* (2.1) *is discrete then the definition can be written replacing integral by sum*

$$\mathcal{F}(f) = F(y) = \sum_{n=-\infty}^{\infty} f(n)e^{-j2\pi ny}, \tag{2.7}$$

where $f(n) : \mathbb{Z} \mapsto \mathbb{C}$.

In this case the image of a function, which is called the Fourier transform of a function, is not the same kind of function than the original one. It results that in the discrete case the Fourier spectrum is again a continuous function. Moreover, the discrete Fourier spectrum is a periodic function.

In a digital device the continuous spectrum cannot be handled. Instead of using a continuous frequency interval only some points of this domain are used. The spectrum definition in (2.7) is evaluated at some frequency points, only. In a practical application only finite many samples are available so the transformation becomes simpler. The name of the resulting transformation is the discrete Fourier transform (DFT).

Definition 2.1.3 (Discrete Fourier transform)**.**

$$F(f_l) = \sum_{k=1}^{N_t} x(t_k) e^{-j2\pi t_k f_l}, \quad l = 1, \ldots, N_f \tag{2.8}$$

where N_t is the number of the samples in time domain, N_f is the number of elements in the frequency grid, f_l is an element from the finite frequency set and t_k is an element from the finite time grid.

The well known expression of DFT is a special case of (2.8). If $x(k)$ denotes $x(t_k)$, $F(f_l)$ is denoted by $F(l)$, $N_f = N_t$, $t_k = kT_s$, $f_l = l\frac{f_s}{N}$, then

$$F(l) = \sum_{k=1}^{N} x(k) e^{-j2\pi kT_s l \frac{f_s}{N}} = \sum_{k=1}^{N} x(k) e^{-j2\pi \frac{kl}{N}} \tag{2.9}$$

using the fact $f_s = \frac{1}{T_s}$.

It is important that sometime only the spectral representation is considered and in this case the Fourier transformation is needless.

In the case of DFT the scalar product is

$$< F, G >= \sum_{l=1}^{N_f} F(f_l) G(f_l) \tag{2.10}$$

where N_f is the number of the frequency points. And the scalar product defines a norm which can be used in the practical cases.

In this section the basic definitions of the Fourier transform were shown. The next section introduces the Hardy spaces which completes the theory of the Fourier transform. The connection between the time domain and the frequency domain is either the Laplace transformation or the z-transformation. The presented definitions of the Fourier transform are special cases of these two transformations but in this thesis only these special cases, in which domains of the transformations (in z-domain the unit circle, in s-domain the imaginary axes) are the stability bounds, are investigated. This is the reason why later s- and z-domain are taken into account.

Here, the bridge between the s-domain and z-domain is the well-known sampling theorem [Schnell, 1993]. The theorem assumes that s-domain signals have finite band width, which means that signals contain no energy above a certain specified maximum frequency. It is called band limited signal assumption which differs from, for example, using a zero-order-hold (ZOH) reconstructor.

2.2 Hardy Spaces

In the previous section the basic Fourier transforms are investigated and elaborated. Now, the results are extended and more properties are shown. The introduction of the so-called Hardy spaces gives insight to classify the different algorithms in the topic of stable approximation. This approach allows to study the space of functions obtained by appropriate Fourier transform and to show the bridge between the s- and z-domain.

Unfortunately, the notations can be different in various articles, books and communities. Here, mainly the notation introduced in [Baratchart et al., 1992], [Zhou et al., 1995] are adopted because of extensive use in the next chapter.

Firstly, an important note from complex analysis is cited [Rudin, 1976].

Note 2.2.1. *[About conformal mapping] Two regions in the complex plane are called conformally equivalent if they are analytically isomorphic, that is, if there exists a one-to-one analytic mapping of the first domain onto the second (whose inverse is then automatically analytic). Using conform functions we can construct a map between $\mathbb{C} \to \mathbb{C}$ where the image set of a circle or a line is again a circle or a line. This property shows us that this map $f(z)$ has the following form:*

$$f(z) = \frac{az + b}{cz + d}. \tag{2.11}$$

Using this we can construct a conformal map between the s-domain and z-domain where the stability bounds are equivalent (Möbius transform): $(z-1)/(z+1)$.

This note says that the z-domain and s-domain are the same from this point of view. Let $\mathbb{D}$ denote the open unit disc in the complex plane $\mathbb{C}$. That is, $\mathbb{D} = \{z \in \mathbb{C}, |z| < 1\}$. Let us denote $\mathbb{U}$ the complement of the closure of the unit disc in the closed complex plane $\overline{\mathbb{C}}$. That is, $\mathbb{U} = \{z \in \overline{\mathbb{C}}, |z| > 1\}$. If f is an analytic function on $\mathbb{D}$, and $r \in (0,1)$, denote $f^{(r)}$ a function on $[0,1]$, defined by $f^{(r)}(\vartheta) = f(re^{j2\pi\vartheta})$. Clearly $f^{(r)}$ is continuous, and periodic. Let

$$\|f^{(r)}\|_2 = \left(\int_0^1 |f^{(r)}(\vartheta)|^2 d\vartheta \right)^{\frac{1}{2}} = \left(\int_0^1 |f(re^{j2\pi\vartheta})|^2 d\vartheta \right)^{\frac{1}{2}}. \tag{2.12}$$

Definition 2.2.1. *$H_2(\mathbb{D})$ (called a Hardy space) denotes the set of all analytic functions f on $\mathbb{D}$ for which*

$$\|f\| = \sup_r \|f^{(r)}\|_2 < \infty. \tag{2.13}$$

$H_2(\mathbb{D})$ is nonempty, as it contains the polynomials. $H_2(\mathbb{D})$ is a Hilbert space, with the inner product

$$< f, g >= \lim_{r \to 1} \int_0^1 f(re^{j2\pi\vartheta})g^*(re^{j2\pi\vartheta})d\vartheta. \tag{2.14}$$

If we write the power series (Taylor) expansion, converging for $|z| < 1$,

$$f(z) = \sum_{k \geq 0} a_k z^k, \tag{2.15}$$

if follows from Parseval's equality that $\|f^{(r)}\|_2^2$ is equal to

$$\sum_{k \geq 0} |a_k|^2 r^{2k}, \tag{2.16}$$

so that f belongs to $H_2(\mathbb{D})$ if and only if

$$\sum_{k \geq 0} |a_k|^2 < \infty. \tag{2.17}$$

Therefore there exists a one-to-one map between the Hardy space $H_2(\mathbb{D})$ and the subspace of $L_2(\mathbb{T})$ consisting of functions whose Fourier coefficients of negative rank are zero. This correspondence associates with f defined by (2.15) the function F in $L_2(\mathbb{T})$ defined by

$$F(e^{j2\pi\vartheta}) = \sum_{k \geq 0} a_k e^{jk2\pi\vartheta}. \tag{2.18}$$

Thus, by definition, the coefficients of the power series expansion of f at zero are the Fourier coefficients of F. F is the natural extension of f to the boundary $\mathbb{D}$, namely the unit circle $\mathbb{T}$. This allows us to consider $H_2(\mathbb{D})$ as a closed subspace of $L_2(\mathbb{T})$.

Like above, the Hardy space $H_2(\mathbb{U})$ can be defined.

Definition 2.2.2. *The Hardy space $H_2(\mathbb{U})$ is the space of functions f, analytic on $\mathbb{U}$, and satisfying*

$$\sup_{r > 1} \left(\int_0^1 |f(re^{j2\pi\vartheta})|^2 d\vartheta \right)^{\frac{1}{2}} < \infty. \tag{2.19}$$

If we write the power series (Laurent) expansion, converging for $|z| > 1$,

$$f(z) = \sum_{k \geq 0} \frac{a_{-k}}{z^k}, \tag{2.20}$$

if follows that the integral in the left hand-side of equation (2.19) is equal to

$$\sum_{k \geq 0} \frac{|a_{-k}|^2}{r^{2k}}, \tag{2.21}$$

so that f belongs to $H_2(\mathbb{U})$ if and only if

$$\sum_{k \geq 0} |a_{-k}|^2 < \infty. \tag{2.22}$$

Therefore there exists a one-to-one map between the Hardy space $H_2(\mathbb{U})$ and the subspace of $L_2(\mathbb{T})$ consisting of functions whose Fourier coefficients of positive rank are zero. This correspondence associates with f defined by (2.20) the function F in $L_2(\mathbb{T})$ defined by

$$F(e^{j2\pi\vartheta}) = \sum_{k \geq 0} a_{-k} e^{-jk2\pi\vartheta}. \tag{2.23}$$

Thus, by definition, the coefficients of the power series expansion of f at infinity are the Fourier coefficients of F. F is the natural extension of f to the boundary $\mathbb{U}$, namely the unit circle $\mathbb{T}$. This allows us to consider $H_2(\mathbb{U})$ as a closed subspace of $L_2(\mathbb{T})$. This again proves that $H_2(\mathbb{U})$ and $H_2(\mathbb{D})$ are Hilbert spaces.

In this thesis only the functions returning real values for real arguments are investigated. They are real subspaces of Hardy spaces $H_2(\mathbb{D})$, $H_2(\mathbb{U})$. It means that they inherit a structure of real Hilbert space. Let us denote the corresponding real subspace of $H_2(\mathbb{U})$ by H_2^+ and the corresponding real subspace of $H_2(\mathbb{D})$ by H_2^-. Moreover, let $L_{2,\mathbb{R}}(\mathbb{T})$ denote the real subspace of $L_2(\mathbb{T})$. Thus the following orthogonal decomposition is true:

$$L_{2,\mathbb{R}}(\mathbb{T}) = H_2^+ \oplus H_2^- \tag{2.24}$$

where $\oplus$ denotes the direct sum. The scalar product on $L_{2,\mathbb{R}}(\mathbb{T})$ is by definition

$$< f, g > = \int_0^1 f(e^{j2\pi t}) g^*(e^{j2\pi t}) dt = \frac{1}{2\pi j} \int_{\mathbb{T}} f(z) g^*(z) \frac{dz}{z}. \tag{2.25}$$

And this induces the Hardy norm on H_2^+ and H_2^-.

There is an another way to introduce the Hardy spaces as subspaces of $L_{2,\mathbb{R}}(\mathbb{T})$. Instead of investigating the Fourier coefficients the time domain representation of signals are used. Without going into details we simply remark that in this case the following subspace decomposition is applied:

$$l_2 = l_2^+ \oplus l_2^-, \tag{2.26}$$

where the space l_2 is defined as

$$l_2 = \left\{ \{a_k\}_{k=-\infty,\dots,\infty} \,\bigg|\, \sum_{k=-\infty}^{\infty} |a_k|^2 < \infty \right\}, \tag{2.27}$$

the subspaces l_2^+ and l_2^- are defined as

$$l_2^+ = \left\{ \{a_k\}_{k=-\infty,\dots,\infty} \in l_2 \,\big|\, a_k = 0 \text{ if } k < 0 \right\}, \tag{2.28}$$

$$l_2^- = \left\{ \{a_k\}_{k=-\infty,\dots,\infty} \in l_2 \,\big|\, a_k = 0 \text{ if } k \geq 0 \right\}. \tag{2.29}$$

The subspace H_2^+ can be identified with the subspace l_2^+ and the subspace H_2^- can be identified with the subspace l_2^- by using the z-transformation.

Between linear subspaces a linear transformation can be defined. If the image of the whole space is a subspace then it is a projection. In Hardy spaces the projections P_+ and P_- can be defined.

$$P_+ : L_2(\mathbb{T}) \mapsto H_2^+,$$
$$P_- : L_2(\mathbb{T}) \mapsto H_2^-. \tag{2.30}$$

For complete description the s-domain case is also introduced where the space L_2 is studied. The projection spaces are $L_2^+ := L_2[0, \infty)$ which is a subspace of L_2 with function zero on the negative real numbers. The projection spaces are $L_2^- := L_2(-\infty, 0]$ which is a subspace of L_2 with function zero on the positive real numbers.

The Hardy space H_2^+ is by definition a closed subspace of $L_2(j\mathbb{R})$ with functions $F(s)$ analytic in $\text{Re}\{s\} > 0$ (open right-half plane). The corresponding norm is defined as

$$\|F\|^2 = \sup_{\sigma > 0} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} F(\sigma + j\omega)F^*(\sigma + j\omega)d\omega \right). \tag{2.31}$$

It can be shown that

$$\|F\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(j\omega)F^*(j\omega)d\omega. \tag{2.32}$$

Similarly, the H_2^- is by definition the orthogonal complement of H_2^+, by name the functions $F(s)$ analytic in $\text{Re}\{s\} < 0$. Figure 2.1 and 2.2 show graphically the relation of different Hardy spaces.
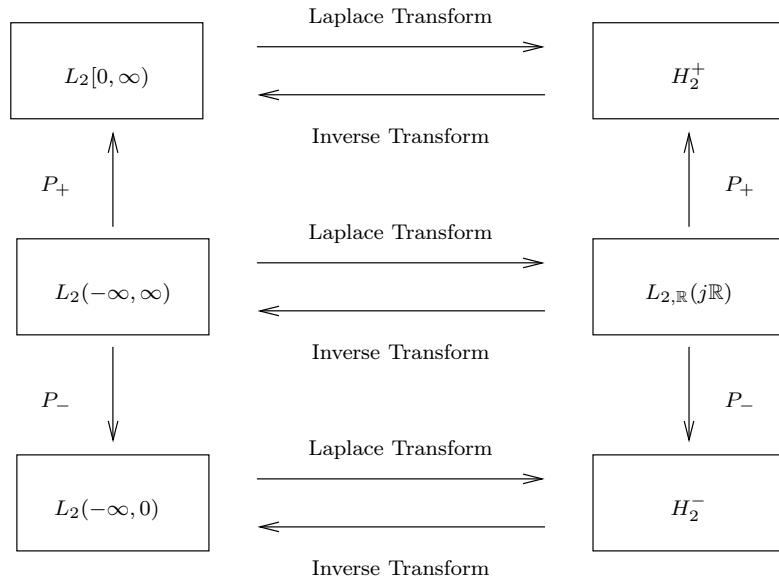
Figure 2.1: Relationship among function spaces in s-domain.
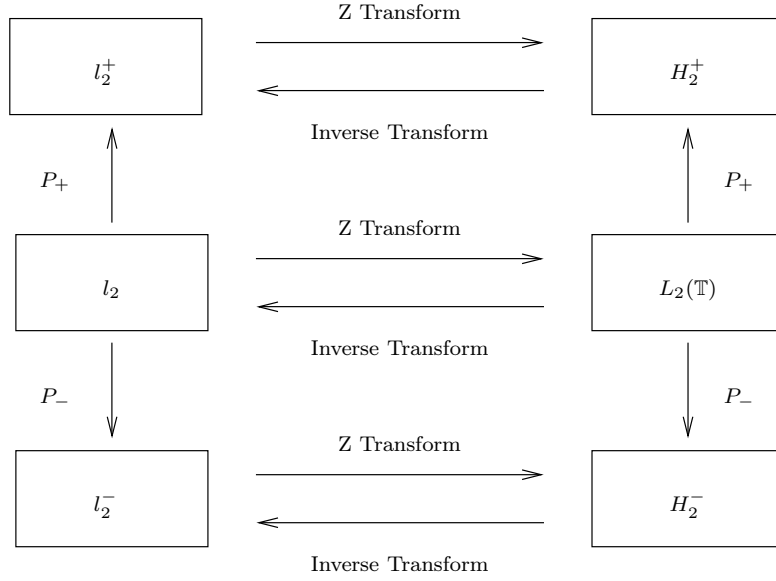
Figure 2.2: Relationship among function spaces in z-domain.

Another important Hardy space in s-domain is the space $L_\infty(j\mathbb{R})$. This is a Banach space, but has no identity element. The norm of the function F

$$\|F\|_\infty = \operatorname*{ess\,sup}_{\omega \in \mathbb{R}}\{F(j\omega)\} \tag{2.33}$$

where ess sup is the essential supremum which is by definition

$$\operatorname*{ess\,sup}_{f(x)} = \inf\{K : \mu(\{x : |f(x)| > K\})\} \tag{2.34}$$

where μ is the measure.

The rational subspace H_∞^+ of L_∞ consists of all rational proper transfer function with no poles on the imaginary axis. H_∞^+ is a closed, real subspace in L_∞ with functions that are analytic in the open right-half plane and bounded on the imaginary axes. The H_∞^+ norm is defined as

$$\|F\|_\infty = \sup_{\mathrm{Re}\{s\}>0}\{F(s)\} = \operatorname*{ess\,sup}_{\omega \in \mathbb{R}}\{F(j\omega)\}. \tag{2.35}$$

The rational subspace H_∞^- of L_∞ consists of all rational proper transfer function with no poles on the imaginary axis. H_∞^- is a closed, real subspace in L_∞ with functions that are analytic in the open right-half plane and bounded on the imaginary axes. The H_∞^- norm is defined as

$$\|F\|_\infty = \sup_{\mathrm{Re}\{s\}<0}\{F(s)\} = \operatorname*{ess\,sup}_{\omega \in \mathbb{R}}\{F(j\omega)\}. \tag{2.36}$$

In the case of discrete time the Hardy space $L_{\infty,\mathbb{R}}(\mathbb{T})$ can be defined similarly.

2.3 Systems and models

In this thesis linear, time invariant systems are studied in continuous ($t = \mathbb{R}$) and discrete ($t = \mathbb{Z}$) cases. Systems are analyzed in the well-known representation: in the frequency domain. In the previous sections the Fourier transformation and the Hardy spaces are investigated. Now, the applied systems are described. Since every linear, time invariant system can be determined by its impulse response, a nice characterization in the frequency domain can be given using tools from the previous subsections. The key of this representation is the Fourier transform which connects the systems represented in time domain with the system represented in frequency domain [Söderström and Stoica, 1989].

The systems presented in this section are elements of the corresponding Hardy spaces. We study functions from the real Hardy subspaces $L_{2,\mathbb{R}}(\mathbb{T})$, H_2^+, H_2^-. The reason is that in the nature only real quantities can be found.

Most of the system in real life is nonlinear and time variant. But in most cases they can be approximated well by linear, time invariant models. Linear, time invariant and continuous time systems can be described by differential equations

$$y(t) + \sum_{r=1}^{n_\beta} \beta_r \frac{d^r y(t)}{dt^r} = \sum_{r=0}^{n_\alpha} \alpha_r \frac{d^r u(t)}{dt^r} \tag{2.37}$$

where $y(t) \in \mathbb{R}$ is the output, $u(t) \in \mathbb{R}$ is the input function, $\beta_r \in \mathbb{R}$ and $\alpha_r \in \mathbb{R}$ are the model parameters. Linear time invariant, discrete time system can be described be difference equations

$$y[t] + \sum_{k=1}^{n_\beta} \beta_k y[t - kT_s] = \sum_{k=0}^{n_\alpha} \alpha_k u[t - kTs], \tag{2.38}$$

where T_s is the sampling time. In both cases system equations (2.37) and (2.38) can be transformed into the following forms (if $n_\alpha \leq n_\beta$):

$$\frac{dx(t)}{dt} = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t) \tag{2.39}$$

or

$$x[t + 1] = Ax[t] + Bu[t]$$
$$y[t] = Cx[t] + Du[t]. \tag{2.40}$$

The name of these representations is state-space variables. $A \in \mathbb{R}^{n \times n}$ is the state matrix, $B \in \mathbb{R}^{n \times 1}$ is the input matrix, $C \in \mathbb{R}^{1 \times n}$ is the output matrix, and $D \in \mathbb{R}^{1 \times 1}$ is the direct term, giving the instantaneous input-output interaction. The state-space representation is not unique. If T is a non-singular matrix, then $T^{-1}AT$, $T^{-1}B$, TC, D matrices describe the same system.

The relationship between the transfer function and the state-space variables can be

obtained by taking the Laplace transform or the z-transform. The transfer function is

$$H(\Omega) = D + C(\Omega I - A)^{-1} B \tag{2.41}$$

where Ω denotes the generalized frequency. In the discrete time case $\omega = [0, 2\pi]$, therefore the function $e^{j\omega}$ maps this interval to the unit circle of the complex plane $\mathbb{C}$. In the continuous case $\omega = \mathbb{R}$ hence the function $j\omega$ maps the whole real axes to the imaginary axes. As a matter of fact this transformation is a rotation on the complex plane. It is very important to note that in both cases the frequency domain is continuous.

The poles of $H(\Omega)$ are the subset of poles of $(\Omega I - A)^{-1}$. The denominator polynomial $L(\Omega)$ of $(\Omega I - A)^{-1}$ can be obtained by using the Cramer's rule for the matrix inverse, i.e.

$$L(\Omega) = \det(\Omega I - A). \tag{2.42}$$

$L(\Omega)$ is the characteristic polynomial of the matrix A. In the case of $\Omega = j\omega$ the set of eigenvalues of A, which are the roots of $L(\lambda)$, can be separated into two parts:

- $\lambda_1, \ldots, \lambda_u$ with real part larger than zero,

- $\lambda_{u+1}, \ldots, \lambda_n$ with real part smaller than zero.

In the case of $\Omega = e^{j\omega}$ the following sets are constituted

- eigenvalues $\lambda_1, \ldots, \lambda_u$ that are outside of the unit disk,

- $\lambda_{u+1}, \ldots, \lambda_n$ that are inside of the unit disk.

The sets may be empty.

Definition 2.3.1. *If there is no eigenvalue with real part larger than zero or there is no eigenvalue outside the unit disk then the corresponding system determined by $G(\Omega)$ is called stable.*

In the case of linear systems this stability coincides with the bounded input-bounded output (BIBO) stability. In the case of BIBO stability we define a stable system to be one for which every bounded input gives rise to an output that also is bounded [Schetzen, 1980].

It can be shown that there exists a non-singular matrix T which separates A in A_u and A_s, respectively the unstable and stable part of A

$$A_{\text{new}} = \begin{bmatrix} A_{u,\text{new}} & 0 \\ 0 & A_{s,\text{new}} \end{bmatrix} = T^{-1} A T. \tag{2.43}$$

Every eigenvalue of $A_{u,\text{new}}$ has positive real part or is outside of the closed unit disk and every eigenvalue of $A_{s,\text{new}}$ has negative real part or inside of the unit disk. $A_{u,\text{new}}$ is called the unstable part of A and $A_{s,\text{new}}$ is called the stable part of A. Similarly, $B_{\text{new}} = T^{-1} B$ and $C_{\text{new}} = CT$ can be obtained. It is known that the system (A, B, C, D) and the system $(A_{\text{new}}, B_{\text{new}}, C_{\text{new}}, D)$ are equivalent, therefore from here we shall drop the subscript "new", and simply write A_u instead of $A_{u,\text{new}}$, and so on.

Using the matrix inversion lemma, it can be shown that $G(\Omega)$ can be decomposed as:

$$H(\Omega) = H_s(\Omega) + H_u(\Omega) + D \tag{2.44}$$

with

$$\begin{aligned} H_s(\Omega) &= C_s(\Omega I_s - A_s)^{-1}B_s \\ H_u(\Omega) &= C_u(\Omega I_u - A_u)^{-1}B_u. \end{aligned} \tag{2.45}$$

We call $H_s(\Omega)$ the stable part of $H(\Omega)$ and similarly $H_u(\Omega)$ the unstable part of $H(\Omega)$.

The parametric model that is used in this thesis is a rational form

$$H(\Omega, \theta) = \frac{N(\Omega, \theta)}{D(\Omega, \theta)} = \frac{\sum_{r=0}^{n_\beta} \beta_r \Omega^r}{\sum_{r=0}^{n_\alpha} \alpha_r \Omega^r} \tag{2.46}$$

where Ω denote the generalized frequency which equals $e^{j\omega}$ in the discrete time case and $j\omega$ in the continuous time case, and with $\theta \in \mathbb{R}^{n_\beta + n_\alpha + 2}$ the vector of the parameters

$$\theta = \begin{bmatrix} \alpha_0 & \alpha_1 & \dots & \alpha_{n_\beta} & \beta_0 & \beta_1 & \dots & \beta_{n_\alpha} \end{bmatrix}. \tag{2.47}$$

In identification problems and approximation problems this parameter vector of the approximator system has to be determined.

It is important to note that all real rational strictly proper transfer functions with no poles on the imaginary axis form a subspace (not closed) of $L_{2,\mathbb{R}}$.

2.3.0.1 Introducing the Delay

In order to extend the model a time delay can be introduced. The new model of the transfer function is

$$H_\tau(\Omega, \theta) = \frac{N(\Omega, \theta)}{D(\Omega, \theta)} e^{-j\omega\tau} \tag{2.48}$$

where τ is the delay value. In the time domain delay is a shifting function along the time axes. In the frequency domain an exponential factor appears in (2.48). Without this factor $H(\Omega, \theta)$ with finite dimension parameter vector in (2.46) can only model the shift along the time axes.

If $H(\Omega, \theta)$ is in $L_{2,\mathbb{R}}$ then $H_\tau(\Omega, \theta) \in L_{2,\mathbb{R}}$.

2.4 Approximation and Parametric Identification

In general, the approximation and the parametric identification are very similar to each other [Boyd and Vandenberghe, 2004]. In both cases an abstract space is given and the measurement data or the target function determines an abstract point from this space. Moreover, a subset (usually a subspace) of points from the space is also given. This subset can be many kinds. In the approximation theory this subset is defined by the

approximator functions. In the identification problems it is defined by the possible models which are parametrised by the so-called parameter vector. The aim is the determination of the nearest point from the defined subset. In both cases the solution of the problem is equivalent with minimization of the so-called cost function. In approximation theory the cost function is usually the appropriate power of the norm. In estimation theory the divergence of distribution functions behaves similarly to a norm.

However, there are differences between the identification and approximation problems. The identification approach means that a parameter, which may be a vector, has to be estimated from the noisy data. The noisy data means that there is a stochastic effect which has influence on the input data. Therefore the estimation problem disregarding the trivial cases results in a stochastic variable. The estimators are characterized by their properties, like for example the variance or the bias of the estimator. In the approximation theory there is no stochastic component. There is a given function on a given domain, and the aim is to find the best approximator where the best means the nearest model in the abstract space. The approximators are characterized by the properties of the error functions. It is important that in identification problems in contrast with approximation problems there are lots of situation where the parameter is a specific physical quantity.

In practise these problems do not arrive independently. There are lots of cases where the results of estimations must be approximated [Vuerinckx, 1998]. For example, the stable approximation, which is one of the topics of this thesis, belongs to this set of problems. If the result of the estimation approach is an unstable model then the approximation of this unstable model by a stable one is significant in some practical situations. An another example is nonlinear system identification. In the nonlinear world the approximation error has more importance than the errors due to the noise. Models which do not describe the nonlinearity in the measurement might lead to unmanageable large errors.

Chapter 3

Stable Identification and/or Approximation

In many cases the model resulting from an identification approach should obey some restrictions. A possible restriction investigated in this thesis is that the model must be a stable one. This constraint can be imposed in an approximation problem, too. The methods in both cases are similar. The measurement data or the target function determines a system and the problem is to find the nearest element to this system from a well defined subset of the whole space.

It sounds as simple as hard to solve. In the case of lots of common problems we known only that the nearest element mentioned above is exist and may be unique, but searching methods are too general or have no effective implementation. Therefore the problems are attacked not generally but individually. This means that there are lots of solved cases both from theoretical and practical points of view. And this means also that there are lots of unsolved cases.

The thesis is about one of the latter case. The stable approximation is a very important question and there are many results with regard to this problem. In this work the problem studied is the stable approximation with additional delay in z- and s-domain. So far the existence of such kind of delay not been proven, for low order systems this thesis fills this gap. Furthermore, in a later chapter, using the connection with the ordinary differential equations, a new numerical method is proposed and analyzed.

This chapter is about earlier published results of the stable identification/approximation. The following chapters contain the new results. In chapter 4 there is a theoretical result and in chapter 5 a new practical algorithm is shown.

3.1 When is a Stable Approximation Needed?

3.1.1 Compensation

The first example for stable approximation is a compensation problem. A possible way to improve the overall quality of mixed analog/digital systems is to introduce a digital compensation for the analog part. In Figure 3.1 two different configurations can be seen.
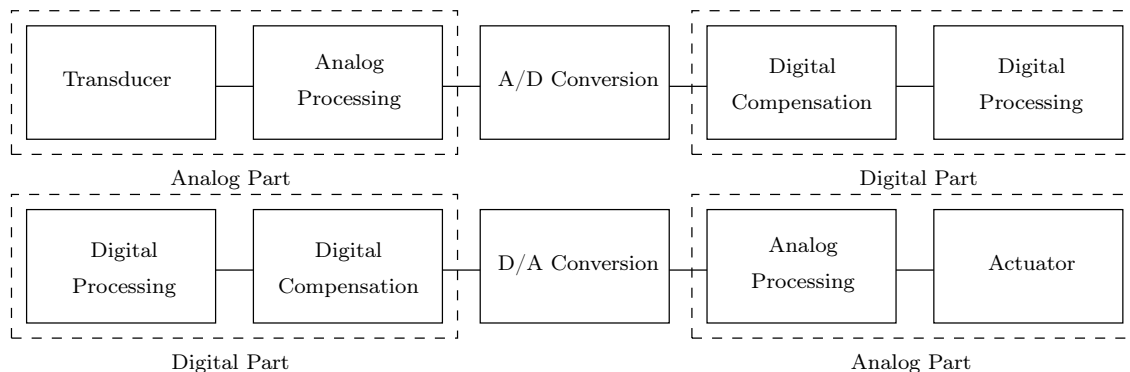
Figure 3.1: Possible configurations for the digital compensation of analog systems.

In the case of upper configuration the transducer must be compensated and the bottom configuration is used when the actuator must be compensated.

The "Analog Processing" block is responsible to comply the Shannon criterion by low-pass filtering. It causes that little aliasing will occur and consequently it is possible to map unambiguously the analog frequency response into the digital world. A possible interpretation of the digital compensation in the actuator case is feed-forward controlling. Set of applications includes the compensation of anti-alias filter in acquisition channel, the compensation of the reconstruction filter in arbitrary waveform generator, loudspeakers compensation, correction of transmission lines, feed-forward control of a laser cutting machine, etc.

There is no theory which guarantees that every function in the world can be approximate completely in the digital world. The definition of every approximation problem wants some restrictions. Even if most analog real-life systems are not linear, nor time invariant, a lot of them can be approximated quite accurately by a linear and time invariant system in their normal operation range. Sometimes a non-linear or a time varying system can be tuned into a linear, time invariant system by using an additional feedback controller. For instance, a transistor is a highly non-linear and rather temperature sensitive device, yet when used with an appropriate feedback circuit, it becomes a nice linear amplifier.

The linear, time invariant system can be described in time or frequency domain. In time domain, it is far from obvious to describe what a good compensation means, in what respect the perfect and the realized overall impulse response may differ, and how this difference will affect the quality of the compensation. In the frequency domain it is quite straightforward because the compensation is not calculated in the whole frequency band, but in subset of all frequencies. Therefore, the digital compensation leads to a digital filter design problem in the frequency domain.

In digital signal processing the linear, time invariant filters can be divided into two sets: FIR filters and IIR filters. Both groups have advantages and disadvantages. The FIR filters possess some good properties, among which are

- A perfect linear phase is possible.

- They are always stable.

- The implementation is easy.

- There exist effective design algorithms.

A consequence of the Weierstrass theorem is that increasing the order of FIR filters can achieve increased approximation accuracy. The advantage of IIR filters is that generally they need less filter orders to achieve the same accuracy. If the purpose of the FIR filter is to compensate an analog system, most of these advantages are irrelevant. In equalization problems a non-linear phase has to be approximated and design of such non-linear phase FIR filters is far from simple although it is still easier than design of corresponding IIR filters.

The problem of a stable approximation is relevant only in the case of IIR filters. In lots of compensation problems a little delay in the digital part can be tolerated, therefore the stable approximation can be achieved by using the techniques presented in this thesis.

3.1.2   Approximation of an Unstable Model After Identification

In figure 3.2 a rough diagram of stable identification can be seen. The identification is started from the noisy data. Even when the measured system is stable, due to the measurement noise and/or nonlinear distortion the resulted model of the identification process might be an unstable one. There are two possible solutions: one-step methods or two-step algorithms. In this chapter existing methods from both sets are introduced.

In the one-step identification method usually the original identification algorithm is modified in order to obtain a stable model. In lots of cases the possible model space is tightened so that in every step the algorithm produces a parameter vector that corresponds a stable model. Here no detail is presented because deeper introductions of some important published methods are in the next sections. The main disadvantage of the classical one-step identification method is that it cannot provide bias bounds, nor accurate uncertainty bounds.

The error bounds can be calculated in the case of two-step procedures. In these methods two independent steps are executed. The first step is a classical unconstrained identification step in which the optimal noise weighting is applied. This step produces the noise error bounds. If the result is a stable model then the identification process ends. If the resulted model of this step is an unstable model then the identification procedure continues with a stable approximation of the unstable model. In this case bias error bounds can be calculated. So at the end of the identification the result is a stable model and the error bounds are available.

The set of stable approximation algorithms can be divided into two subsets:

- No delay is added to the target function.

- Some delay is introduced to the target function.

Both cases are introduced here. This thesis covers the case of stable approximation with additional delay.
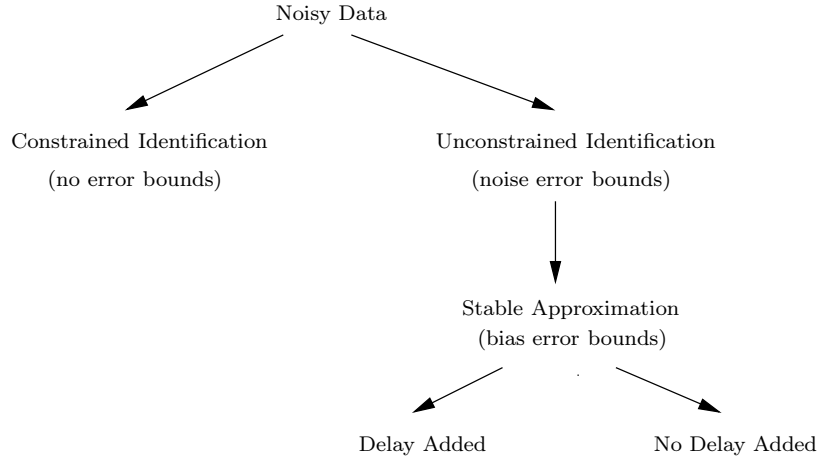
Figure 3.2: One-step vs. two-step stable identification.

The overview of existing methods is in the next two sections. First, the constrained identification methods are investigated and then some of the published post-processing algorithms are shown.

3.1.2.1 System Inversion

In control theory, system inversion approaches have been already investigated. In [Silverman, 1969] and [Sain and Massey, 1969] constructive methods for the inversion of multi-variable systems are discussed. The idea behind the so-called structure algorithm is to incorporate delayed measurements in the output equation. The system $\tilde{G}(z)$ is an L-delay inverse for $G(z)$ if

$$\tilde{G}(z)G(z) = \frac{1}{z^L}I_m$$

where I_m is the identity matrix and m is the number of inputs. The resulting inverse system then consists of delay elements followed by the inverse of the transformed system. Sain and Massey also considered inverses which consist of a bank of delay elements followed by a dynamical system. In their approach, however, the number of differentiators needed to realize the inverse can be higher than the order of the original system. An important concept, introduced by Sain and Massey, is the inherent delay of a discrete-time system, which is the minimal delay with which the input of the system can be reconstructed from the system outputs. The inversion procedure of Sain and Massey allows to derive inverses with arbitrary delay larger than or equal to the inherent delay. A disadvantage of the inversion procedure described above is that it can yield unstable inverses.

3.2 Constrained Stable Identification/Approximation

3.2.1 Subspace Methods

In subspace identification the system matrices are estimated based on estimated Kalman filter state sequences and the observed inputs and outputs. For a finite number of data points, the estimated system matrix is not guaranteed to be stable. The stability can be imposed by for example data augmentation [Chui and Maciejowski, 1996] or using regularization [Van Gestel et al., 2001].

The linear stochastic identification problem which contains the deterministic problems as a subset is concerned with systems and models of the form

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + w_k \\
y_k &= Cx_k + Du_k + v_k
\end{aligned}
\tag{3.1}
$$

with

$$
\mathbb{E}\left(\begin{bmatrix} w_p \\ v_p \end{bmatrix} \begin{bmatrix} w_p & v_p \end{bmatrix}\right) = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \geq 0
\tag{3.2}
$$

where $\mathbb{E}$ denotes the expected value operator. The subscript k is the discrete time and vectors $u_k \in \mathbb{R}^m$ and $y_k \in \mathbb{R}^l$ denote the m inputs and l outputs of the system respectively. The number of states is n and $x_k \in \mathbb{R}$ is the state vector at the time index k of the system with order n. The process noise $w_k \in \mathbb{R}^n$ and the measurement noise $v_k \in \mathbb{R}^l$ are assumed to be zero mean, Gaussian with covariance matrix as given by (3.1). The model matrices A, B, C, D and the covariance matrices Q, S, R have appropriate dimensions.

In the last decade, so-called subspace identification methods [Van Overschee and De Moor, 1996], [Katayama, 2005] have been developed to determine the system order $\hat{n}$ of the unknown system (3.1) and the estimates $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$ (up to a similarity transformation) together with the estimated noise covariance matrix $\hat{Q}$, $\hat{S}$, $\hat{R}$, from a large number of observations of the input u_k and the corresponding output y_k generated by the unknown system. The identification can be divided into two steps:

- Estimation of the state sequence $\hat{X}_i, \hat{X}_{i+1} \in \mathbb{R}^{\hat{n} \times j}$ by Kalman filter using geometric operations.

- Estimation of the model matrices $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$ by minimizing the following cost function

$$
\min_{\hat{A},\hat{B},\hat{C},\hat{D}} \left\| \begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} - \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} \begin{bmatrix} \hat{X}_i \\ U_{i|i} \end{bmatrix} \right\|_F^2
\tag{3.3}
$$

where $U_{i|i} = [u_i, u_{i+1}, \ldots, u_{i+j-1}] \in \mathbb{R}^{m \times j}$ and $Y_{i|i} = [y_i, y_{i+1}, \ldots, y_{l+j-1}] \in \mathbb{R}^{l \times j}$.

It is known that the estimation in the subspace method is consistent as $j \to \infty$.

Generally, if θ denotes the parameter vector then the regularization is obtained by adding a so-called regularization term, which may be $c\|\theta\|_2$ or $c\|\theta\|_2^2$, to the cost function

$J(\theta)$. Thus the new cost function is $J(\theta) + c\|\theta\|_2$ or $J(\theta) + c\|\theta\|_2^2$ [Boyd and Vandenberghe, 2004]. In the basic form of regularized approximation, the goal is to find a vector θ that is small (if possible), and also makes the cost function $J(\theta)$ small. In this case a specific type of the regularization term is used. Let W be positive-semidefinite matrix and $W = W^T$. Then the regularization term added to the cost function (3.3) is

$$\text{trace}\{\hat{A}W\hat{A}^T\}. \tag{3.4}$$

Adding the regularization term to the cost function (3.3) a stable model may be obtained.

It is possible to include the minimum phase constraint in the subspace identification, see details in [Tanaka and Katayama, 2005].

3.2.2   Stable Approximation in Hardy Spaces

There is an intensive research on the approximation theory in the different Hardy spaces [Baratchart et al., 1992], [Baratchart et al., 1996], [Baratchart et al., 1997]. In these papers the approximation problems are investigated in mathematical sense. Lots of theories on the existence and uniqueness can be found there. It is important to note that the problems described in these papers are finding the best rational approximation from the stable subset of the rational functions. Hence, the stability is guaranteed in the formulation of the problems. In this subsection some results with respecting these papers and properties of described approximators are presented.

Minimization in Space H_2

The question studied in [Baratchart et al., 1992] is the following one: given an integer n and some $f \in H_2^-$, find a minimum of the squared norm

$$\left\| f - \frac{p}{q} \right\|^2, \tag{3.5}$$

where p/q ranges over rational functions in H_2^-, subject to the constraint that degree of q is less than n. The function $\Phi : (p, q) \mapsto \mathbb{R}$ is defined as

$$\Phi : (p, q) \mapsto \left\| f - \frac{p}{q} \right\|^2 \tag{3.6}$$

where p belongs to the set of polynomials of degree at most $n-1$, denoted by $R[z]_{n-1}$, and q belongs to the set of monic polynomials of degree n whose roots are inside the unit disk, denoted by $R[z]_n^-$. It can be seen that Φ is not injective because the possible cancellation of common factors of q and p. However, the following lemma says that this will happen only in trivial case.

Lemma 3.2.1. *If $f \in H_2^-$ is not a rational function of degree strictly less than n, then the argument of any local minimum of* (3.5) *is an irreducible fraction whose degree is exactly n.*

It proves that the local minima of (3.5) coincide with local minima of Φ. Φ is a differentiable function and its critical points (zeros of its gradient field) are investigated. The set $R[z]_{n-1} \times R[z]_n^-$ can be identified with an open set of $\mathbb{R}^{2n}$ in the following way:

$$p(z) = p_0 + p_1 z + \cdots p_{n-1} z^{n-1},$$
$$q(z) = q_0 + q_1 z + \cdots q_{n-1} z^{n-1} + z^n. \tag{3.7}$$

is identified with the points $(p_{n-1}, \ldots, p_0) \in \mathbb{R}^n$ and $(q_{n-1}, \ldots, q_0) \in \mathbb{R}^n$. Let us denote V_q the n dimensional subspace of H_2^- generated by

$$\frac{1}{q}, \frac{z}{q}, \frac{z^2}{q} \cdots , \frac{z^{n-1}}{q}. \tag{3.8}$$

We have that p/q is the projection of f onto V_q. This means that the search for critical points can be restricted to pairs (p, q). In this way, p becomes a function of q denoted by $L(q)$. Let us define $\psi_N(q)$ as

$$\psi_n(q) : \mathbb{R}[z]_n^- \mapsto \mathbb{R} = q \mapsto \left\| f - \frac{L(q)}{q} \right\|^2. \tag{3.9}$$

Lemma 3.2.2. *The map ψ_n is a smooth function. Its critical points are the same as those of Φ.*

The existence of a minimum can be proven in a more general case. Another important question is uniqueness. It has been proven that uniqueness of the best approximant is strongly generic property, that is, true on an open dense subset of H_2^-. However, there are situations where more than one absolute minima of the cost function exist. Therefore additional conditions are required to ensure the uniqueness. In has turned out that a possible condition is if the function f is analytical in a wider domain. Such condition prevents the function f having a sequence of zeros accumulating on the unit circle. Hereafter we shall assume that f is holomorphic in an open disk U_r, of radius r, with $r > 1$.

Let Δ_n be the closure in $\mathbb{R}^n$ of the set $\mathbb{R}[z]_n^-$. Δ_n is a compact set which consists of monic polynomials of degree n whose roots are of modulus less or equal to 1. Denote P_r the open neighborhood of Δ_n consisting of monic polynomials whose roots are of modulus strictly less than r.

Lemma 3.2.3. *If f is holomorphic in U_r then Ψ_n extends to a smooth function*

$$\Psi_n : P_r \mapsto \mathbb{R}. \tag{3.10}$$

Lemma 3.2.4. *The set Δ_n is homeomorphic to the closed unit ball B^n of $\mathbb{R}^n$.*

Therefore the boundary $\partial \Delta_n$ of Δ_n is homeomorphic to the $(n-1)$ dimensional sphere S_{n-1}. Example Δ_2 which is a triangle shows that the it is not smooth. The smooth part of $\partial \Delta_n$ is the polynomials having exactly one irreducible factor over $\mathbb{R}$. The following

functions are embeddings

$$\phi_1 : \dot{\Delta}_{n-1} \mapsto \Delta_n \qquad \phi_1(p) = (z-1)p$$
$$\phi_{-1} : \dot{\Delta}_{n-1} \mapsto \Delta_n \qquad \phi_{-1}(p) = (z+1)p$$
$$\phi_c : (-1,+1) \times \dot{\Delta}_{n-2} \mapsto \Delta_n \quad \phi_c(\alpha, p) = (z^2 + 2\alpha z + 1)p \qquad (3.11)$$

where $\dot{\Delta}_n$ is the interior Δ_n. Thus the smooth part of $\partial \Delta_n$ is the union of the images ϕ_1, ϕ_{-1} and ϕ_c.

The following three lemmas give the basic of the minimum seeking algorithm published. The proof of the lemmas and a detailed description of this algorithm can be found in [Baratchart et al., 1992].

Lemma 3.2.5. *Let $q \in \partial \Delta_n$ be such that $q = \chi p.$, where all the roots of χ are of modulus 1, and $p \in \dot{\Delta}_k$ where the degree of the polynomial p is k. Then $L(q) = \chi L(p)$ and $\phi_n(q) = \phi_k(p)$.*

Lemma 3.2.6. *Let $p \in \dot{\Delta}_{n-1}$, and $x = \phi_1(p)$, or $\phi_{-1}(p)$. The projection of $\nabla \phi_n(x)$ on $\partial \Delta_n$ coincides with $\nabla \phi_{n-1}(p)$. If $x = \phi_c(\alpha, p)$, $\alpha \in (-1, 1)$, then the projection of $\nabla \phi_n(x)$ on tangent space $T_x \partial \Delta_n$ lies in the subspace $\phi_c(\alpha, \Delta_{n-2})$, where it coincides with $\nabla \phi_{n-2}(p)$.*

Lemma 3.2.7. *Let p be a minimum of ϕ_{n-1}, and $x = \phi_1(p)$, or $\phi_{-1}(p)$ or else $\phi_c(\alpha, p)$ with $\alpha \in (-1, 1)$, then $\nabla \phi_n(x)$ is orthogonal to $T_x \partial \delta_n$ and points outwards (if non zero).*

And finally the algorithm is the following.

1. Choose an initial point q_0.

2. Integrate the vector field $-\Delta \phi_n$ from the initial conditions $(q_0, \phi_n(q_0))$.

   - If a local minimum is reached, then end.
   - If the boundary $\partial \Delta_n$ is reached, then goto 2.

3. You are at the point q_b of $\partial \Delta_n$:

$$q_b = \chi q_i, \chi(\alpha) = 0 \Rightarrow |\alpha| = 1, q_i \in \dot{\Delta}_k. \qquad (3.12)$$

   Integrate the vector field $-\nabla \phi_k$ from the initial conditions $(q_i, \phi_k(q_i))$.

   - If the minimum at order $k < n$ is reached, the goto 3.
   - If the boundary of Δ_k is reached, replace n by k and goto 2.

4. You are at a minimum $q_m \in \dot{\Delta}_k$ of ϕ_k. Integrate the vector field $-\nabla \phi_{k+1}$ from the initial conditions $((z+1)q_m, \phi_k(q_m))$.

   - The minimum at order $k+1$ is reached. If $k+1 < n$ the replace k by $k+1$ and goto 3. If $k+1 = n$ then end.
   - $\partial \Delta_{k+1}$ is reached. Replace n by $k+1$ and goto 2.

The algorithm does not stop on a saddle point because it is an unstable critical point. Let us define the function $\phi(q) = \phi_k(p)$. At each step this function decreases hence the same minimum cannot occur twice.

The result of this algorithm is not necessary a global minimum but a local one. Without any more assumption we cannot guarantee that it is global.

Minimization in Space H_∞

The presented results can be found in [Baratchart et al., 1996]. Let K be a subset of the unit circle in the complex plane, i.e. $K \subset \mathbb{T}$. In this article the following problems are addressed.

Let us given $f \in L_\infty(K)$, $h \in L_\infty(\mathbb{T}\backslash K)$, and $M > 0$. We want to find $g \in H_\infty$ whose distance to h in $L_\infty(\mathbb{T}\backslash K)$ does not exceed M and which is as close as possible to f in the $L_\infty(K)$ metric under this constraint.

If $K = \mathbb{T}$ then the problem is equivalent with the standard Nehari problem (see [Zhou et al., 1995]). The precise formulation of the problem addressed in the article [Baratchart et al., 1996]:

Problem 3.2.1. *Let K be a subset of $\mathbb{T}$ of positive measure. For $f \in L_\infty(K)$ and $h \in L_\infty(\mathbb{T}\backslash K)$, define a subset $\mathcal{B}_{M,h}$ of H_∞ by:*

$$\mathcal{B}_{M,h} := \{g \in H_\infty, \|h - g\|_{L_\infty(\mathbb{T}\backslash K)} \leq M\}, \tag{3.13}$$

where M is a positive real number. We seek $g_0 \in \mathcal{B}_{M,h}$ such that

$$\|f - g_0\|_{L_\infty(K)} = \min_{g \in \mathcal{B}_{M,h}} \|f - g\|_{L_\infty(K)}. \tag{3.14}$$

Let $H_\infty|_K$ denote the restriction of H_∞ onto K.

Theorem 3.2.1. *Let K be a subset of $\mathbb{T}$ such that $\lambda(K) > 0$. Then:*

- *$H_\infty|_K$ is not dense in $L_\infty(K)$.*

- *If K is open, the closure of $H_\infty|_K$ is contained in $H_\infty|_K + C(K)$.*

- *If K is a proper closed subset of $\mathbb{T}$, then $\mathcal{A}|_K$ is dense in $C(K)$ and the closure of $H_\infty|_K$ in $L_\infty(K)$ contains $(H_\infty + C(\mathbb{T}))|_K$.*

In the sense of approximation an important proposition of this theorem is the following.

Proposition 3.2.1. *Let K be a subset of $\mathbb{T}$ such that $\lambda(K) > 0$, and f be in $L_\infty(K)$. Suppose (g_n) is a sequence of H_∞ functions such that $g_n|_K$ converges to f in $L_\infty(K)$. If f is not the trace of an H_∞ function, then*

$$\lim_{n \to \infty} \|g_n\|_{L_\infty(\mathbb{T}\backslash K)} = \infty. \tag{3.15}$$

3.2.3   Stable Identification with ELiS

In the Frequency Domain System Identification (fdident) toolbox an iterative, gradient-type method is implemented to minimize a cost function [Pintelon and Schoukens, 2001], [Kollár, 1994], [Kollár, 1993]. The result is the maximum likelihood estimation of the parameter vector in the frequency domain. The name ELiS refers to the estimation of linear systems. The estimated system is described in the §2.3.

Detailed description about the assumptions and methods in frequency domain identification can be found in §6.2. In chapter 6 we do not care whether the results of estimators are stable or not. But now, we concentrate only on the stability, therefore just a quick and minimal overview is given.

This paragraph of the thesis covers an additional topic of the stable approximation in frequency domain, too. As already mentioned in §2.4, approximation and identification are closely related to each other. Because of the historical order results are introduced here and not in §3.3. As it will be seen, via a little modification of the cost function of ELiS the approximation can be formulated in the frequency domain. It was already mentioned in [Pintelon and Schoukens, 1990] and in [D'haene et al., 2006] the methods are summarized and extended.

We use an error-in-variables framework and, therefore, the direct division of the measured input and output is not used. The noise free inputs and outputs are $U_0(\Omega_k)$ and $Y_0(\Omega_k)$, $k = 1, \ldots, F$, where F is the number of frequencies. The input and output spectra may be computed by DFT or may be measured directly.

Every input $U_0(\Omega_k)$ and every output $Y_0(\Omega_k)$ are distorted by additional noise processes which are denoted by $N_U(\Omega_k)$ and $N_Y(\Omega_k)$. The relationships between the measured and the noiseless variables are

$$U(\Omega_k) = U_0(\Omega_k) + N_U(\Omega_k)$$
$$Y(\Omega_k) = Y_0(\Omega_k) + N_Y(\Omega_k). \tag{3.16}$$

The maximum likelihood cost function whose global minimum we seek is [Pintelon et al., 1994], [Pintelon and Schoukens, 2001]

$$C_{\mathrm{ML}}(\theta) = \sum_{k=1}^{F} \frac{|N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)|^2}{\sigma_{U,k}^2 |N(\Omega_k, \theta)|^2 + \sigma_{Y,k}^2 |D(\Omega_k, \theta)|^2 - 2\mathrm{Re}\{\sigma_{YU,k}D(\Omega_k, \theta)N^*(\Omega_k, \theta)\}} \tag{3.17}$$

where F is the number of the frequencies, $\sigma_{U,k}^2$ is the variance of the $N_U(\Omega_k)$, $\sigma_{Y,k}^2$ is the variance of the $N_Y(\Omega_k)$ and $\sigma_{YU,k}^2$ is the covariance between $N_Y(\Omega_k)$ and $N_U(\Omega_k)$. Because of the whitening property of the DFT their distributions are assumed to be independent Gaussian with zero mean and corresponding variances $\sigma_{U,k}$, $\sigma_{Y,k}$.

It can be seen that this minimization is nonlinear in θ. Methods that minimize this kind of cost functions are described in §5.5. The algorithms are sequences of points in the parameter space. At every point calculation of the next step is based on the gradient function. It is worth noting that the variance of the numerator coincides with the denomi-

nator of (3.17). Due to the measurement noise and/or the nonlinear distortions the result of the estimation may be an unstable model. To overcome this problem the steps of the gradient-like minimization must be modified.

In approximation problems there are no variance/covariance variables and there are no input and output. Now, if we impose for all k, $U(\Omega_k) = 1$, $\sigma_{U,k} = 0$, $\sigma_{Y,k} = 1$ and $Y(\Omega_k) = T(\Omega_k)$ where $T(\Omega_k)$ is desired transfer function, the minimization of (3.17) is equivalent to solving the least squares approximation problem. It enables us to discuss the two problems simultaneously.

We will use the notation introduced in §2.3. Let $H(\Omega, \theta)$ denote the transfer function in the frequency domain. $H(\Omega, \theta)$ can be written as sum of the three terms (see (2.44)).

- Leaving out the unstable part $H_u(\Omega, \theta)$. This is the optimal approximant in L_2 norm but results a lower order model. However, in bandpass approximation the optimality is no longer true.

- Reflection: in this case stability is enforced by reflecting the unstable poles with respect to the theoretical stability bound after each iteration step. The reflection does not change the amplitude response, but it does change the phase response. By this, the phase of the fitted model decreases, and so for a proper fit a negative delay (non-causal system) would be adequate. This step is nothing else than replacing the

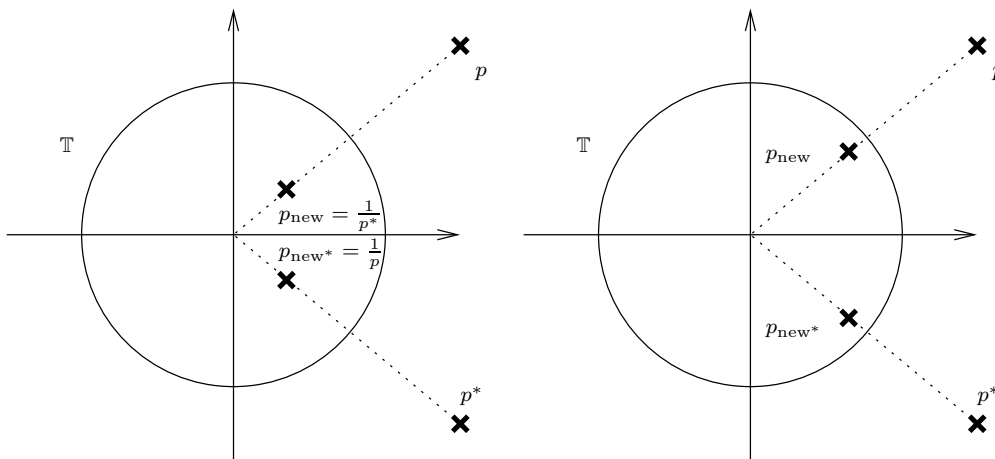Figure 3.3: An example of inverting the poles (left) and an example of contracting the poles (right) in z-domain. $\mathbb{T}$ is the unit circle.

unstable matrix A_u part by A_u^{-1}. In the z-domain every pole outside the unit disk is reflected with respect to the unit circle. In s-domain all the unstable poles are reflected with respect to the imaginary axes. The positions of the zeros remain the same.

- Replacing $H_u(s,\theta)$ by $H_u(-s,\theta)$ in s-domain or replacing $H_u(z,\theta)$ by $H_u(z^{-1},\theta)$ in z-domain result the following new transfer functions

$$\hat{H}(s,\theta) = H_u(-s,\theta) + H_s(s,\theta) + D$$
$$\hat{H}(z,\theta) = H_u(z^{-1},\theta) + H_s(z,\theta) + D. \tag{3.18}$$

This changes the position of the poles and zeros of $H_u(\Omega,\theta)$ instead of only the poles of $H_u(\Omega,\theta)$.

- Contraction: stability is enforced by contraction of the unstable poles to the stability limit after each iteration step. Let $p = r_p e^{j\theta_p} = x_p + jy_p$ denote an unstable pole of $H(\Omega,\theta)$. Contraction is replacing p by $-\varepsilon + jy_p$ in s-domain or replacing p by $(1-\varepsilon)e^{j\theta_p}$ in z-domain, where ε is a given small number.

The methods above are used to generate initial values for the non-linear minimization algorithm. A gradient-based algorithm can improve the model. In order to preserve the stability the non-linear minimization must be modified. As in §5.5 one can see one step of the gradient type algorithm is the update of the parameter vector:

$$\theta_{k+1} = \theta_k + \Delta\theta. \tag{3.19}$$

Knowing that θ_k corresponds with a stable model, there are two possible results. Either θ_{k+1} determines a stable model, and one can start with the next iteration, or θ_{k+1} determines an unstable model, and $\Delta\theta$ is reduced, for example, a factor 2. This procedure results a stable model at least as good as the initial values.

3.3 Stable Approximation as a Post-Processing Step

In this section two step identification procedures introduced in §3.1.2 are investigated. As already mentioned, we distinguish two different approaches: adding delay to the target function or not. Modification of the target function with additional delay cannot always be allowed. In a closed-loop application it is not possible. But in an open-loop case where an additional delay does not cause any problem, it is a very effective method in order to obtain a stable model.

3.3.1 The Closest Stable Polynomial

The method described previously in this section defines the best stable approximation in the space of poles and zeros. However, the approximation is usually done in the space of coefficients. In [Moses and Liu, 1991] a method which uses the coefficient space is presented. The paper considers solutions to the stabilization problem that minimizes the error between the polynomial coefficients of the original and the stabilized polynomial. Since the polynomial coefficients are being estimated, it is natural to stabilize the polynomial by perturbing these estimated coefficients as little as possible.

The norm is an L_2 norm. Let $b = [b_1, \ldots, b_n]^T$ be a real vector and let $B(z)$ denote the associated monic polynomial in z-domain:

$$B(z) = z^n + b_1 z^{n-1} + \cdots + b_{n-1} z + b_n. \tag{3.20}$$

Our assumption is that $B(z)$ is unstable, i.e. has at least one zero z_0 satisfying $|z_0| > 1$. The problem is to find a vector $a = [a_1, \ldots, a_n]$ such that the measure of distance J is minimal and the polynomial associated to a is stable. Let the measure of distance be defined as

$$J = \|a - b\|_W^2 = (a - b)^T W (a - b) \tag{3.21}$$

for some given positive definite weighting matrix W. Let us define the following set

$$S_a = \{a | A(z) = 0 \Rightarrow |z| < 1\}. \tag{3.22}$$

Using this definition the problem can be formulated as

Problem 3.3.1. *Given a vector $b \notin S_a$, find a vector $a^o \in S_a$ such that J is minimized over all $a \in S_a$.*

Using the Schur parameters (also known as reflection coefficients) the set S_a can be completely parameterized by S_r which is a subset of the space of Schur parameters. The main advantage of this parametrization is that S_r is a hypercube, thus a convex set.

The advantage of this method is that it is easy to implement. On the other hand, it is not guaranteed that the optimal stable minimum of the original approximation problem coincides with the result of this method. In some special cases the weighting matrix W may be computed using the original cost function, but it is not necessary.

3.3.2 Stable Approximation with Addition Delay

In this section a new method to obtain a stable model is introduced. The investigation and analysis are the main topics of the next two chapters of this thesis. The main difference between this and the previous approach is that a new variable, the delay is introduced in the cost function. The advantage is that the global minimum of the cost function at an appropriate delay value determines a stable model. The main disadvantage is that the result of this method can be used only in an open loop configuration.

For example, simply inverting the transfer function of an analog system without additional delay usually leads to a bad approximation. This is because most analog systems that must be equalized have a non-minimum phase model in the z-plane, which implies that they introduce some delay that cannot be compensated for with a causal filter.

3.3.2.1 Problem Formulation

Let Ω denote the generalized frequency (for s-domain $\Omega = j\omega$ and for z-domain $\Omega = e^{j\omega}$). The so-called original stable approximation problem is the following. Minimize the

following cost function:

$$\int_I |T(\Omega) - H(\Omega, \theta)|^2 d\omega \text{ w.r.t } \theta, \tag{3.23}$$

where $T(\Omega)$ is the target system, $H(\Omega, \theta)$ is the approximator

$$H(\Omega, \theta) = \frac{\beta_0 + \beta_1\Omega + \ldots + \beta_{n_\beta}\Omega^{n_\beta}}{\alpha_0 + \alpha_1\Omega + \ldots + \alpha_{n_\alpha}\Omega^{n_\alpha}} \tag{3.24}$$

with n_β and n_α is the order of numerator and denominator of $H(\Omega, \theta)$, respectively, θ is the parameter vector

$$\theta = \begin{bmatrix} \beta_0 & \ldots & \beta_{n_\beta} & \alpha_0 & \ldots & \alpha_{n_\alpha} \end{bmatrix}. \tag{3.25}$$

I denotes the domain of integration. It can be a full domain (for s-domain $I = (-\infty, \infty)$ and for z-domain $I = (0, 2\pi)$) or a subset of the whole frequency band.

This kind of problems were introduced and investigated already in this section. This is an approximation problem in a Hardy space. Until this point only methods without additional delay were considered. The main aim is to minimize this cost function so that the parameter vector is stable. Now the cost function is rewritten in order that the global minimum determines a stable model. The new cost function is

$$C(\theta, \tau) = \int_I |T(\Omega)e^{-j\omega\tau} - H(\Omega, \theta)|^2 d\omega, \tag{3.26}$$

where τ is the delay.

Since $T(\Omega)e^{-j\omega\tau}$ is still in $L_{2,\mathbb{R}}(\mathbb{T})$ or in $L_{2,\mathbb{R}}(j\mathbb{R})$, the extended approximation problem can be formulated using Hardy spaces.

Increasing the delay value causes that the global minimum of the cost function (3.26) determines a stable model. So the approximation without imposing any constraints results a stable model $H(\Omega, \theta)$.

Minimization of the cost function (3.26) is in the focus of the thesis, therefore detailed introductions are in chapter 4 and in 5.

3.3.2.2   About Norms

The problem (3.26) is an optimization problem in an L_2 space. In practise other norms have importance, too. Since the distance between two elements in a normed space is defined similarly, the cost function practically remains the same. Apart from the notation identity the problems defined by different normed spaces can differ significantly from each other. For example, applying the Chebyshev norm (namely L_∞) causes that the cost function is no more a continuous function with respect to the parameter vector. In [Vuerinckx, 1998] some numerical experience shows that the minima of cost functions defined by the norm L_2 and L_∞ are close to each other. Unfortunately, this experience has no theoretical background. There is a result in [Ellacott and Williams, 1976b] that says that the global

minimum of the cost function defined by L_∞ norm can be computed by doing optimization in the space L_2 using iteratively modified weighting, [Ellacott and Williams, 1976a].

3.3.2.3   Results of the Thesis

Since the result of the thesis belongs to the topic of stable approximation with additional delay in the target function, a quick overview is given here.

As mentioned in [Vuerinckx, 1998] adding a delay to the target function in (3.26) causes that the global minimum yields a stable model. In the literature only intuitive proofs are published. It is one of the main result of this thesis that a mathematically correct proof is presented in the next chapter (chapter 4).

We will prove that for all $T(\Omega)$ (fixed during minimization) with $n_\beta + n_\alpha \leq 2$ there exists a delay τ such that the global minimum θ_* of the cost function (3.26) yields a stable model $H(\Omega, \theta_*)$.

Furthermore, some consequences of this theorem are investigated.

Having studied theoretical results a new numerical method is presented. This method is able to find a stable model which is the minimum of the cost function and the corresponding delay value automatically.

Chapter 4

Existence Proof

4.1   Problem description

Let Ω denote the generalized frequency (for s-domain $\Omega = j\omega$ and for z-domain $\Omega = e^{j\omega}$). Minimize the following cost function (τ is fixed):

$$\int_I |T(\Omega)e^{-j\omega\tau} - H(\Omega,\theta)|^2 d\omega \text{ w.r.t } \theta, \tag{4.1}$$

where T is the target system

$$T(\Omega) = \frac{\gamma_0 + \gamma_1\Omega + \cdots + \gamma_{n_{\beta T}}\Omega^{n_{\beta T}}}{\delta_0 + \delta_1\Omega + \cdots + \delta_{n_{\alpha T}}\Omega^{n_{\alpha T}}} = c_T \frac{(\Omega - z_{T,1})\cdots(\Omega - z_{T,n_{\beta T}})}{(\Omega - p_{T,1})\cdots(\Omega - p_{T,n_{\alpha T}})} \tag{4.2}$$

with $n_{\beta T}$ and $n_{\alpha T}$, respectively, the order of the numerator and the denominator of $T(\Omega)$; $H(\Omega,\theta)$ is the approximator

$$H(\Omega,\theta) = \frac{\beta_0 + \beta_1\Omega + \ldots + \beta_{n_\beta}\Omega^{n_\beta}}{\alpha_0 + \alpha_1\Omega + \ldots + \alpha_{n_\alpha}\Omega^{n_\alpha}} = c_H \frac{(\Omega - z_1)\cdots(\Omega - z_{n_\beta})}{(\Omega - p_1)\cdots(\Omega - p_{n_\alpha})} \tag{4.3}$$

with n_β and n_α, respectively the order of numerator and denominator of $H(\Omega,\theta)$, θ is the parameter vector

$$\theta = \begin{bmatrix} \beta_0 & \ldots & \beta_{n_\beta} & \alpha_0 & \ldots & \alpha_{n_\alpha} \end{bmatrix} \tag{4.4}$$

and τ is the delay. I denotes the domain of integration. It can be a full domain (for s-domain $I = (-\infty, \infty)$ and for z-domain $I = (0, 2\pi)$) or a subset of the whole frequency band.

We introduce notations for the numerator and the denominator of $H(\Omega,\theta)$:

$$N(\Omega,\theta) = \beta_0 + \beta_1\Omega + \ldots + \beta_{n_\beta}\Omega^{n_\beta} \tag{4.5}$$

and

$$D(\Omega,\theta) = \alpha_0 + \alpha_1\Omega + \ldots + \alpha_{n_\alpha}\Omega^{n_\alpha}. \tag{4.6}$$

33

Hence, we can write that

$$H(\Omega, \theta) = \frac{N(\Omega, \theta)}{D(\Omega, \theta)}. \tag{4.7}$$

The dimension of the parameter space is $n_\beta + n_\alpha + 2$. Unfortunately the map from the parameter space to the space of transfer functions is onto but not one-to-one. To overcome this problem the dimension of the parameter vector space must be decreased. It can be done by fixing one coefficient of the numerator or the denominator (projection onto a subspace) or by fixing the norm of the parameter vector θ (projection onto the unit sphere). This reduction of dimension provides that the global minimums can be characterized unambiguously [Pintelon et al., 1994].

We will prove that for all $T(\Omega)$ (fixed during minimization) and for $n_\beta + n_\alpha \leq 2$ there exists delay τ such that the global minimum θ_* of the cost function (4.1) determines a stable model $H(\Omega, \theta_*)$. The generalized case for arbitrary order will be handled with presuming two reasonable assumptions.

4.2 Notes

Before starting the proof some preliminary notes are shown in order to simplify the proof.

Note 4.2.1. $X, Y \in \mathbb{C}$, $X \neq 0$, $X + Y \neq 0$:

$$\frac{1}{X + Y} = \frac{1}{X}\left(1 - \frac{Y}{(X + Y)}\right) = \frac{1}{X} - \frac{Y}{X(X + Y)} \tag{4.8}$$

The so-called Landau notations are introduced. In this thesis only infinite asymptotics are used.

Definition 4.2.1. $O(.)$ *estimate:* $f(x)$ *and* $g(x)$ *are functions defined for all* x *sufficiently large.* $f(x) \in O(g(x))$ *means that there exists constants* x_0 *and* c *such that* $|f(x)| \leq c|g(x)|$ *for all* $x \geq x_0$.

Definition 4.2.2. $o(.)$ *estimate:* $f(x)$ *and* $g(x)$ *are functions defined for all* x *sufficiently large.* $f(x) \in o(g(x))$ *means that* $g(x) \neq 0$ *for sufficiently large* x *and* $\lim_{x \to \infty} f(x)/g(x) = 0$.

In both cases, i.e. $f(x) \in O(g(x))$ and $f(x) \in o(g(x))$ we will use the notations:

$$f(x) = O(g(x)), \quad f(x) = o(g(x)). \tag{4.9}$$

(This is a slight abuse of notation; equality of two functions is not asserted, and it cannot be since the properties of being $O(g(x))$ and $o(g(x))$ are not symmetric.)

Definition 4.2.3. *Asymptotic equivalence:* $f(x) \sim g(x)$ *means that* $g(x) \neq 0$ *for sufficiently large* x *and* $\lim_{x \to \infty} f(x)/g(x) = 1$.

Note 4.2.2. *In order to simplify the calculation with the Landau notations their basic properties are listed.*

1. If $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$ then $f_1(x)f_2(x) \in O(g_1(x)g_2(x))$.

2. $f(x) \cdot O(g(x)) \in O(f(x)g(x))$.

3. If $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$ then $f_1(x) + f_2(x) \in O(g_1(x) + g_2(x))$. This implies that $f_1(x), f_2(x) \in O(g(x)) \Rightarrow f_1(x) + f_2(x) \in O(g(x))$.

4. $f(x) + O(g(x)) \in O(f(x) + g(x))$.

5. Multiplication by a constant: let $k \in \mathbb{R}$ and $k > 0$. $O(k \cdot g(x)) = O(g(x))$. $f(x) \in O(g(x)) \Rightarrow k \cdot f(x) \in O(g(x))$.

6. $o(f(x)) + o(f(x)) \subset o(f(x))$.

7. $o(f(x))o(g(x)) \subset o(f(x)g(x))$.

8. $o(o(f(x))) \subset o(f(x))$.

9. $o(f(x)) \subset O(f(x))$. And thus the last three properties apply with most combinations of $o(.)$ and $O(.)$.

Note 4.2.3. *In this note a_n and b_n denote sequences as $n \to \infty$.*

1. If $a_n = o(1)$ then

$$\frac{1}{1 - a_n} = 1 + a_n + O(a_n^2) = 1 + a_n + o(a_n). \tag{4.10}$$

   Specially

$$\frac{1}{1 + O(a_n)} = 1 + O(a_n). \tag{4.11}$$

2. Using note 4.2.1 we have

$$\frac{1}{a_n + o(a_n)} = \frac{1}{a_n(1 + o(1))} = \frac{1}{a_n} - \frac{1}{a_n}\frac{o(1)}{(1 + o(1))} = \frac{1}{a_n} + o\left(\frac{1}{a_n}\right).$$

3. From the previous item

$$\frac{b_n + o(b_n)}{a_n + o(a_n)} = \frac{b_n + o(b_n)}{a_n} + (b_n + o(b_n))o\left(\frac{1}{a_n}\right) = \frac{b_n}{a_n} + o\left(\frac{b_n}{a_n}\right)$$

Note 4.2.4. *$T(z)$ is analytic in an neighborhood of P and $a_n = o(1)$. Then*

$$T(P + a_n) = T(P) + T'(P)a_n + O(a_n^2) = T(P) + O(a_n). \tag{4.12}$$

Note 4.2.5. *Let a_n and b_n be series such that $a_n > 0$, $b_n > 0$, $a_n \to 0$, $b_n \to 0$ and*

$$\frac{a_n}{b_n} \to \infty \quad (n \to \infty) \tag{4.13}$$

then

$$\forall c > 0 : \exists N : \quad n > N \Rightarrow a_n > cb_n. \tag{4.14}$$

Proof. It is a rewritten version of the definition of convergence to infinity. □

Note 4.2.6.

$$\left(1 - \frac{c}{n}\right)^n \to e^{-c} \text{ as } n \to \infty \tag{4.15}$$

Proof. Elementary analysis. □

Note 4.2.7. *If $h_n \to 0$ then*

$$\left(1 + \frac{h_n}{n}\right)^n \to 1 \text{ as } n \to \infty \tag{4.16}$$

Proof. See analysis books. □

Note 4.2.8. *Let us assume that the sequence $d_\tau > 0$, $\forall \tau : d_\tau < \tau$, $d_\tau \overset{\tau \to \infty}{\to} \infty$, $\frac{d_\tau}{\tau} \overset{\tau \to \infty}{\to} 0$. Then for all sufficiently large τ*

$$\left(1 - \frac{d_\tau}{\tau}\right)^\tau < K^{d_\tau} \tag{4.17}$$

where $0 < K < 1$ an appropriate constant.

Proof.

$$\left(1 - \frac{d_\tau}{\tau}\right)^\tau = \left(\left(1 - \frac{d_\tau}{\tau}\right)^{\frac{\tau}{d_\tau}}\right)^{d_\tau} = \left(e^{-1} + o(1)\right)^{d_\tau}. \tag{4.18}$$

By the definition of convergent sequences $\exists D : \forall \tau > D : e^{-1} + o(1) < e^{-1} + 10^{-3} < 1$. Therefore choosing $K = e^{-1} + 10^{-3}$ proves the note. □

Note 4.2.9. *For each sequence d_τ which converges to the positive infinity the following is true. If $|K| < 1$ and $L \in \mathbb{Z}_+$ then*

$$d_\tau^L K^{d_\tau} = o(1). \tag{4.19}$$

Note 4.2.10. *Let us assume that for all n: $h_n > 0$ and $g_n > 0$. Then*

$$\frac{g_n}{h_n^2} \to \infty \quad \Rightarrow \quad \frac{\sqrt{g_n}}{h_n} \to \infty \tag{4.20}$$

or equivalently

$$h_n^2 = O(g_n) \quad \Rightarrow \quad h_n = O(\sqrt{g_n}). \tag{4.21}$$

Definition 4.2.4. *$T(z)$ is analytic in an neighborhood of z_0. $[T(z)]_{z=z_0}^{(m-1)}$ is defined as the $(m-1)$th derivative of $T(z)$ w.r.t. z at $z = z_0$.*

Note 4.2.11. *$\forall \varepsilon > 0 \exists \tau > 0 : \forall p \in \mathbb{C}(|p| \leq 1) \exists q \in \mathbb{C}$ such that $|p - q| < \varepsilon$, $|p| = |q|$ and*

$$\cos(\tau \angle q) = s_1 \cos(\tau \angle p)$$

$$\sin(\tau \angle q) = s_2 \sin(\tau \angle p) \tag{4.22}$$

where $s_1, s_2 \in \{-1, 1\}$ are given.

Proof. It is enough to prove the statement for $|p| = 1$. Let $p = e^{j\alpha}$ where α is the angle of p. We have the following constraints

$$|\sin(\tau\alpha)| = |\sin(\tau\alpha + \Delta\alpha)|,$$

$$|\cos(\tau\alpha)| = |\cos(\tau\alpha + \Delta\alpha)|. \tag{4.23}$$

$\Delta\alpha = k\pi$ in order to fulfil (4.22). Therefore

$$\left|\frac{\Delta\alpha}{\tau}\right| = \frac{|k|\pi}{\tau} \tag{4.24}$$

and

$$q = \exp\left(j\alpha + j\frac{\Delta\alpha}{\tau}\right). \tag{4.25}$$

Moreover,

$$|p - q| = \left|e^{j\alpha} - e^{j\alpha}e^{j\frac{\Delta\alpha}{\tau}}\right| = \left|1 - e^{j\frac{\Delta\alpha}{\tau}}\right|$$

$$= 2\left|\sin\left(\frac{\Delta\alpha}{2\tau}\right)\right| \leq \frac{|\Delta\alpha|}{\tau}. \tag{4.26}$$

Hence for all τ sufficiently large

$$\left|\frac{\Delta\alpha}{\tau}\right| < \varepsilon. \tag{4.27}$$

$\square$

In the proof of the previous note it can be noted that

$$|p| = |q| + O\left(\tau^{-1}\right). \tag{4.28}$$

Intuitively this note says that if we increase τ then the function z^τ changes the signs of its real and imaginary parts more frequently. See figures 4.1 and 4.2.
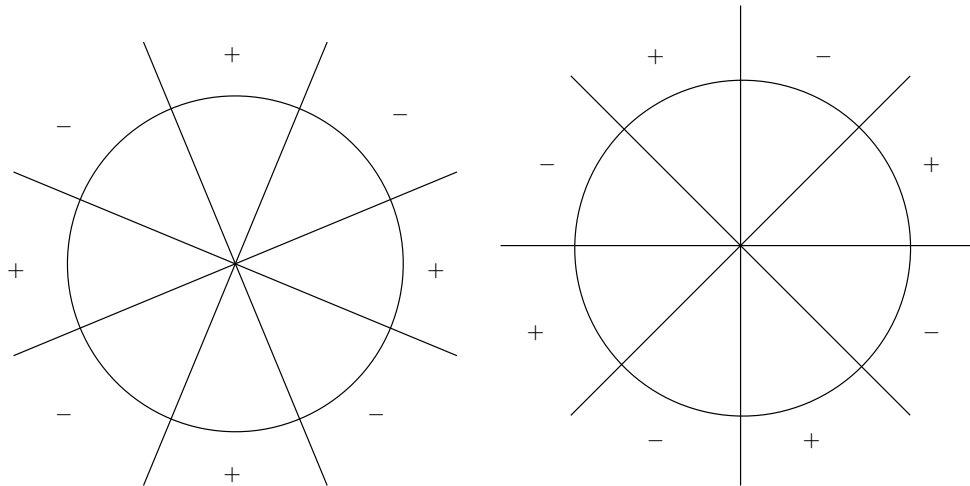
Figure 4.1: Illustration of the sign of $\mathrm{Re}\{z^\tau\}$ (left) and $\mathrm{Im}\{z^\tau\}$ (right) when $\tau = 4$.
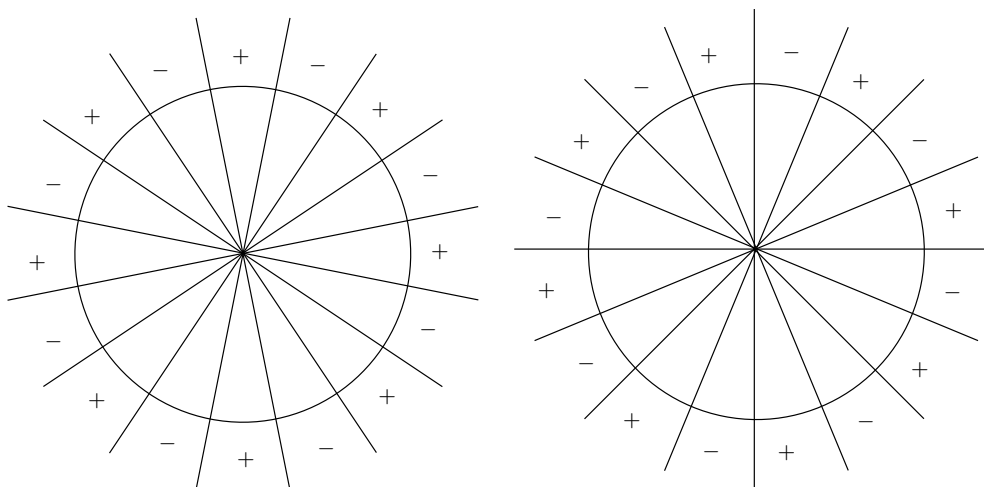
Figure 4.2: Illustration of the sign of $\mathrm{Re}\{z^\tau\}$ (left) and $\mathrm{Im}\{z^\tau\}$ (right) when $\tau = 8$.

A very important consequence is the following note:

Note 4.2.12. *Let $H(z)$ be a meromorphic function which has no pole on the unit circle $\mathbb{T}$.*

$$\forall \varepsilon > 0, \exists \tau > 0 : \left[\forall p \in \gamma_U, \forall s_1, s_2 \in \{-1, 1\} \exists q \in \mathbb{T} : |p - q| < \frac{2\pi}{\tau} \Rightarrow \right.$$

$$\left| H^{(k)}(p) - H^{(k)}(q) \right| < \varepsilon \text{ and } \cos(\tau \angle q) = s_1 \cos(\tau \angle p)$$

$$\text{and } \sin(\tau \angle q) = s_2 \sin(\tau \angle p)] \tag{4.29}$$

where $k \leq n_\beta + n_\alpha + 2$ and $H^{(k)}(z)$ is the kth order derivative of $H(z)$ w.r.t. z.

Proof. While there is no pole on the unit circle and the number of poles are limited, there exists a neighborhood of the unit circle such that on the closure of the neighborhood $\exists M > 0 \forall k : |H^{(k)}| < M$, and all these functions are uniformly continuous. This fact and the previous note proves the statement. $\square$

Definition 4.2.5 (Laurent expansion and residue). *If $f(z)$ is an analytic function on the annulus $R_1 \leq |z - z_0| \leq R_2$ with $0 \leq R_1 \leq R_2 \leq \infty$, then*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n(z - z_0)^n \tag{4.30}$$

for all z in the annulus. a_{-1} is called the residue of f at z_0, denoted $\text{Res}_{z=z_0} f(z)$

$$\text{Res}_{z=z_0} f(z) = \frac{1}{(m-1)!}[(z - z_0)^m f(z)]_{z=z_0}^{(m-1)} \tag{4.31}$$

where m is the order of the pole z_0 $(a_n = 0$ for $n < -m)$.

Definition 4.2.6 (Index of a closed curve). *The index of a closed curve γ $(s \notin \gamma)$ is*

$$n(\gamma, s) = \frac{1}{2\pi i} \int_\gamma \frac{dz}{z - s}. \tag{4.32}$$

For example; let s be an arbitrary point inside the unit circle and let γ be the unit circle (positive direction and one turn):

$$\int_{|z|=1} \frac{dz}{z} = \int_0^{2\pi} \frac{z'(t)}{z(t)} dt = \int_0^{2\pi} \frac{ie^{it}}{e^{it}} dt = 2\pi i \tag{4.33}$$

Note 4.2.13. *Residue theorem (simplified version): Let's γ be a closed curve for which $n(\gamma, s) = 0$ $(s \notin D)$ (like unit circle) and let $f(z)$ regular in the domain D except the isolated singularity. Then*

$$\frac{1}{2\pi i} \int_\gamma f(z) dz = \sum_j n(\gamma, z_j) \text{Res}_{z=z_j} f(z). \tag{4.34}$$

Note 4.2.14. *Complex mean-value theorem: Suppose Ω is an open convex set in $\mathbb{C}$, suppose f is a holomorphic function $f : \Omega \to \mathbb{C}$, and suppose a, b are distinct points in Ω. Then there exist points u, v on L_{ab} (the straight line connecting a and b not containing the endpoints), such that*

$$\text{Re}\left\{\frac{f(b) - f(a)}{b - a}\right\} = \text{Re}\left\{f'(u)\right\},$$

$$\text{Im}\left\{\frac{f(b)-f(a)}{b-a}\right\} = \text{Im}\left\{f'(v)\right\}. \tag{4.35}$$

Definition 4.2.7 (Divided difference). *For distinct real numbers $x_1, x_2, \ldots, x_n$, the divided difference of the function $f\mathbb{R} \mapsto \mathbb{R}$ is defined as*

$$f[x_1] = f(x_1) \tag{4.36}$$

and

$$f[x_1, x_2, \ldots, x_n] = \frac{f[x_1, x_2, \ldots, x_{n-1}] - f[x_2, x_3, \ldots, x_n]}{x_1 - x_n} \tag{4.37}$$

for all $n \geq 2$.

Note 4.2.15. *The n-point divided difference of f can be expressed as*

$$f[x_0, x_1, \ldots, x_n] = \sum_{k=1}^{n} \frac{f(x_k)}{\displaystyle\prod_{\substack{l=1 \\ l \neq k}}^{n}(x_l - x_k)} \tag{4.38}$$

Proof. See Theorem 2.10 in [Sahoo and Riedel, 1998]. $\qquad\square$

Note 4.2.16. *In [Sahoo and Riedel, 1998] it was shown that $f[x_0, x_1, \ldots, x_n]$ is a continuous function of the variables $x_0, x_1, \ldots, x_n$. If $f(x)$ has a continuous nth derivative, then an unique continuous extension of $f[x_0, x_1, \ldots, x_n]$ can be presented by an integral representation. For example, if $n = 1$, then the continuous extension of $f[x_0, x_1]$ is*

$$f[x_0, x_1] = \begin{cases} \dfrac{f(x_0) - f(x_1)}{x_0 - x_1} \text{ if } x_1 \neq x_0, \\ f'(x_0) \text{ if } x_1 = x_0 \end{cases} \tag{4.39}$$

provided $f(x)$ has the first derivative. Because of this unique extension it can be allowed that some of the nodes to coalesce if f is suitable differentiable.

Note 4.2.17. *Mean value theorem for divided differences: Let $f : [a,b] \mapsto \mathbb{R}$ be a real valued function with continuous nth derivative and $x_0, x_1, \ldots, x_n$ in $[a, b]$. Then there exists a point η in the interval*

$$[\min\{x_0, x_1, \ldots, x_n\}, \max\{x_0, x_1, \ldots, x_n\}] \tag{4.40}$$

such that

$$f[x_0, x_1, \ldots, x_n] = \frac{f^{(n)}(\eta)}{n!}. \tag{4.41}$$

Proof. See Theorem 2.10 in [Sahoo and Riedel, 1998]. $\qquad\square$

4.3 The Statement

The note 2.2.1 says that it is enough to prove the existence theorem only in either of the z-domain or s-domain.

Theorem 4.3.1. *For every meromorphic transfer function $T(z) \in L_2(\mathbb{T})$ and for $n_\beta + n_\alpha \leq 2$ there exists a delay τ such that the global minimum θ_* of the cost function with fixed τ below gives a stable system $H(z, \theta_*)$.*

$$C(\theta, \tau) = \parallel T(z)z^{-\tau} - H(z, \theta) \parallel_2^2, \tag{4.42}$$

where θ is the parameter vector that contains the coefficients of the numerator and denominator, respectively.

4.4 The Proof

It is enough to prove the theorem for integer delay values, i.e. we can assume that $\tau \in \mathbb{Z}$.

We can write the cost function into the following form:

$$C(\theta, \tau) = T_1 + T_2(\theta, \tau) + T_3(\theta) \tag{4.43}$$

where

$$T_1 \quad = \quad \int_I |T(\Omega)|^2 d\omega$$

$$T_2(\theta, \tau) \quad = \quad -2\mathrm{Re} \cdot \left\{ \int_I T^*(\Omega) e^{i\tau\omega} H(\Omega, \theta) d\omega \right\}$$

$$T_3(\theta) \quad = \quad \int_I |H(\Omega, \theta)|^2 d\omega \tag{4.44}$$

where star denotes the complex conjugate and $I = [0, 2\pi]$.

The first term in the cost function is independent of θ, and hence we do not care about it furthermore.

The second term in the cost function is responsible for decreasing the cost function. Since the sign of this term can be chosen arbitrarily by the sign of c_H in (4.3), it really reduces the cost function if it is not zero. Let us denote the second term as $T_2(\theta, \tau)$.

The third term is independent of the delay and has always a non-negative value. During the minimization of the cost function it should be reduced as much as possible. Let us denote the third term as $T_3(\Omega, \theta)$.

Let us study the second term in more details. Using $\Omega = e^{i\omega}$ we can write it in the following form:

$$-\frac{T_2(\theta, \tau)}{2} = \mathrm{Re} \left\{ \int_I T^*(\Omega) e^{i\omega\tau} H(\Omega, \theta) d\omega \right\}$$

$$= \text{Re} \left\{ \int_{[0,2\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega \right\} \tag{4.45}$$

It can easily be proved that the operation Re{} is unnecessary:

$$\int_{[0,2\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega$$

$$= \int_{[0,\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega + \int_{[\pi,2\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega \tag{4.46}$$

and using the substitution $\omega_1 = -\omega + 2\pi$ in the second term, we find

$$\int_{[\pi,2\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega = \int_{[\pi,0]} T(e^{i\omega_1})e^{-i\tau\omega_1}H(e^{=i\omega_1},\theta)d(-\omega_1)$$

$$= \int_{[0,\pi]} T(e^{i\omega_1})e^{-i\tau\omega_1}H(e^{-i\omega_1},\theta)d\omega_1. \tag{4.47}$$

Therefore, (4.46) can be elaborated as

$$= \int_{[0,\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega + \int_{[0,\pi]} T(e^{i\omega_1})e^{-i\tau\omega_1}H(e^{-i\omega_1},\theta)d\omega_1$$

$$= \text{Re} \left\{ \int_{[0,\pi]} T(e^{-i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega \right\}. \tag{4.48}$$

Let τ be a positive integer value. (For non-integer values the power function is defined by the complex logarithm function which is regular on the cut plane only. But if $\tau \in \mathbb{N}$ then the function z^τ is regular on the whole $\mathbb{C}$.) Substituting of $e^{i\omega}$ by z gives

$$z = e^{i\omega} \quad \frac{dz}{d\omega} = ie^{i\omega} = iz \tag{4.49}$$

$$\int_{[0,2\pi]} T^*(e^{i\omega})e^{i\tau\omega}H(e^{i\omega},\theta)d\omega = -i\int_{\mathbb{T}} T(z^{-1})z^{\tau-1}H(z,\theta)dz \tag{4.50}$$

where $\mathbb{T}$ is the unit circle. And using note 4.2.13 we get

$$-\frac{T_2(\theta,\tau)}{2} = 2\pi \sum_k \text{Res}_{z=p_k} T(z^{-1})z^{\tau-1}H(z,\theta). \tag{4.51}$$

For $\tau > 1$ the terms in the sum correspond to the poles of either the target or the approximator function. Obviously we are talking about poles in the interior of the unit disk, i.e. the stable poles of $H(z,\theta)$ and the unstable ones of $T(z^{-1})$.

Following the same lines for the third term $T_3(\theta)$ of the cost function we find

$$T_3(\theta) = \int_I |H(\Omega, \theta)|^2 d\omega = \int_{[0,2\pi]} H^*(e^{i\omega}, \theta) H(e^{i\omega}, \theta) d\omega$$

$$= -i \int_{\mathbb{T}} H\left(z^{-1}, \theta\right) z^{-1} H(z, \theta) dz$$

$$= 2\pi \sum_k \text{Res}_{z=p_k} H\left(z^{-1}, \theta\right) z^{-1} H(z, \theta). \tag{4.52}$$

The poles of the system $H(z^{-1}, \theta)$ are specified by inverting the poles of the system $H(z, \theta)$ with respect to the unit circle. Using equation (4.3), equation (4.52) can be elaborated as:

$$= 2\pi \sum_k \text{Res}_{z=p_k} c_H^2 z^{n_\alpha - n_\beta - 1} \frac{(z - z_1)(1 - zz_1) \cdots (z - z_{n_\beta})(1 - zz_{n_\beta})}{(z - p_1)(1 - zp_1) \cdots (z - p_{n_\alpha})(1 - zp_{n_\alpha})}. \tag{4.53}$$

In equation (4.52) the number of terms is $\max\{n_\beta - n_\alpha + 1, 0\} + n_\alpha$ where the term $\max\{n_\beta - n_\alpha + 1, 0\}$ stems from the term $z^{n_\alpha - n_\beta - 1}$ in (4.53). Note that $T_3(\theta) = 0$ if and only if $c_H = 0$.

In the following let θ_* or $\theta_*(\tau)$ denote the series of the parameter vector which for every τ minimize the cost function.

Note that if $\theta = \theta_*$ in (4.51) there is always at least one term. If this was not true, i.e. every pole of $T(z^{-1})$ and $H(z, \theta_*)$ is outside of the unit circle, then inverting a pole or a pair of poles of $H(z, \theta_*)$ would increase $|T_2(\theta_*, \tau)|$ and leave $T_3(\theta_*)$ unchanged and hence would decrease the cost function. But this conflicts with θ_* being a global minimum.

It is important to note that

$$|T_2(\theta_*, \tau)| > |T_3(\theta_*)| \tag{4.54}$$

and thus with an appropriate choice of c_H

$$T_2(\theta_*, \tau) + T_3(\theta_*) < 0. \tag{4.55}$$

If this was not true then $T_2(\theta_*, \tau) + T_3(\theta_*)$ is always larger than zero and the parameter vector with zero c_H would give a smaller cost function ($T_2(\theta, \tau) + T_3(\theta) = 0$) than $C(\theta_*, \tau)$ which is a contradiction. The practical consequence of (4.54) is that the larger $|T_2(\theta, \tau)|$ the smaller the cost function (4.42).

In the remaining part of the proof we use the notation $p(\tau)$ for a pole and $z(\tau)$ for a zero in order to emphasize the dependency on the delay τ. If we need more than one pole or zero then indexing is used. For example $p_1(\tau)$, $z_E(\tau)$, etc. Later, the index is used to distinguish between subsets of the poles and zeros. Simultaneously $p_1(\tau)$ denotes a particular sequence of a pole. Fortunately this does not cause any problem because in every case the correct meaning is given. In the same way τ' denotes one integer value or

a sub-sequence of the positive numbers $\mathbb{N}$ at the same time.

The next steps of the proof are:

1. $T_2(\theta_*, \tau)$ is bounded and therefore so is $T_3(\theta_*)$.

2. There is at least one pole $p(\tau)$ of $H(z, \theta_*)$ which converges to the unit circle, i.e. $\exists p(\tau) : |p(\tau)| \to 1$ as $\tau \to \infty$.

3. If θ_* is a global minimum then $|p(\tau)|$ cannot converge to 1 faster than $o\left(\frac{1}{\tau}\right)$. And there is at least one pole $p_1(\tau)$ for which $|1 - p_1(\tau)| \sim \frac{c_1}{\tau}$ where $c_1 > 0$ is constant.

4. Proof for the case $n_\beta = 1$ and $n_\alpha = 1$.

5. Proof for the case $n_\beta = 0, n_\alpha = 2$ and both poles are complex.

6. Proof for the case $n_\beta = 0, n_\alpha = 2$ and both poles are real.

4.4.1 Evaluating the Term $T_2(\theta, \tau)$

We prove this step by contradiction. Let us assume that there exists a sub-sequence τ' of τ such that

$$T_2(\theta_*(\tau'), \tau') \to \infty \text{ as } \tau' \to \infty. \tag{4.56}$$

First we prove that in this case $\| H(\Omega, \theta_*(\tau')) \|_2 \to \infty$ as $\tau' \to \infty$.

$$|T_2(\theta_*(\tau'), \tau')| = 2 \left| \int_I T^*(\Omega) e^{i\tau'\omega} H(\Omega, \theta_*(\tau')) d\omega \right| \tag{4.57}$$

Using the Cauchy-Schwarz inequality

$$\left| \int_I T^*(\Omega) e^{i\tau'\omega} H(\Omega, \theta_*(\tau')) d\omega \right| \leq \sqrt{\int_I |T(\Omega)|^2 d\omega} \sqrt{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega}$$

$$= \| T(\Omega) \|_2 \cdot \| H(\Omega, \theta_*(\tau')) \|_2 \tag{4.58}$$

Since $\| T(\Omega) \|_2$ is constant as a function of τ', (4.56) and (4.58) prove the statement.

$$\frac{|T_3(\theta_*(\tau'))|}{|T_2(\theta_*(\tau'), \tau')|} = \frac{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega}{2 \left| \int_I T^*(\Omega) e^{i\tau'\omega} H(\Omega, \theta_*(\tau')) d\omega \right|} \tag{4.59}$$

Using the same estimation

$$\geq C_1 \frac{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega}{\sqrt{\int_I |T(\Omega)|^2 d\omega} \sqrt{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega}} \geq C_2 \sqrt{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega} \tag{4.60}$$

where C_1, C_2 are constants, and where

$$\sqrt{\int_I |H(\Omega, \theta_*(\tau'))|^2 d\omega} = \| H(\Omega, \theta_*(\tau')) \|_2 \to \infty \tag{4.61}$$

Hence

$$\frac{|T_3(\theta_*(\tau'))|}{|T_2(\theta_*(\tau'), \tau')|} \geq \| H(\Omega, \theta_*(\tau')) \|_2 \to \infty \tag{4.62}$$

Since this is in contradiction with (4.54), and therefore $|T_2(\theta_*(\tau), \tau)| < \infty$ and $|T_3(\theta_*(\tau))| < \infty$.

4.4.2   Evaluating the Factor c_H

Let $\widehat{H}(\Omega, \theta)$ be defined as

$$\widehat{H}(\Omega, \theta) = \frac{H(\Omega, \theta)}{c_H} = \frac{(z - z_1) \cdots (z - z_{n_\beta})}{(z - p_1) \cdots (z - p_{n_\alpha})}. \tag{4.63}$$

This is slight abuse of notation because $\widehat{H}(\Omega, \theta)$ does not depend on the parameter c_H. However, as it is mentioned in the very first part of this chapter it is nothing else than a non-linear map which reduces the dimension of the parameter space, which makes it possible to determine the local minimums unambiguously. In the remaining part of the existence proof it will be unambiguous when the notation $\widehat{H}(z, \theta)$ is used. Since

$$T_2(\theta, \tau) = -2c_H \int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega$$

$$T_3(\theta) = c_H^2 \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega \tag{4.64}$$

the cost function (4.42) can be written as

$$C(\theta, \tau) = c_H^2 Q(\theta) - 2c_H R(\theta, \tau) + W. \tag{4.65}$$

The same abuse of notations is true for $Q(\theta)$ and $R(\theta, \tau)$, like for $\widehat{H}(\Omega, \theta)$: in the cases of $Q(\theta)$ and $R(\theta, \tau)$ θ is the parameter vector which does not contain c_H. (4.65) can be minimized w.r.t. c_H:

$$= Q(\theta) \left(c_H^2 - 2c_H \frac{R(\theta, \tau)}{Q(\theta)} \right) + W = Q(\theta) \left(c_H - \frac{R(\theta, \tau)}{Q(\theta)} \right)^2 - \frac{R^2(\theta, \tau)}{Q(\theta)} + W. \tag{4.66}$$

The minimum can be reached if c_H is chosen as

$$c_H = \frac{R(\theta, \tau)}{Q(\theta)} \tag{4.67}$$

and the corresponding value of the cost function is

$$C(\theta, \tau) = W - \frac{R^2(\theta, \tau)}{Q(\theta)} = \int_I |T(\Omega)|^2 d\omega - \frac{\left(\int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega\right)^2}{\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega}. \qquad (4.68)$$

It is worth noting that minimization of the cost function (4.42) is equivalent to maximizing the following expression:

$$\overline{C}(\theta, \tau) = \frac{\left(\int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega\right)^2}{\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega}. \qquad (4.69)$$

4.4.3 Subsequences of Poles and Zeros

Throughout the proof, subsequences of poles and zeros will be examined. This part of the thesis presents a tool which simplifies the forthcoming steps of the proof. The basic idea is the following. Assuming that the optimal parameter vector has unstable poles, a new parameter vector is constructed such that it defines only stable poles and gives a lower cost function. Since the optimal parameter vector minimizes the cost function, this is a contradiction and, hence, for all τ sufficiently large the optimal parameter vector defines only stable poles. To apply this tool, in which unstable poles are replaced by stable ones, it is necessary to exclude special cases in which the path of a pole or a zero behaves irregularly. For example, a sequence can have more than one limit point. Here, the correct mathematical background is given to avoid repeating always how the special cases are handled. The presented tool enables us to consider a convergent sequence of poles or zeros. It is worth noting that the applied topological space is the extended complex plane $\mathbb{C} \cup \{\infty\}$ [Rudin, 1987].

The sequences of zeros are defined as the sequences of roots of the following polynomial defined by $\theta(\tau)$, $\tau \in \mathbb{Z}$, $\tau \to \infty$:

$$\beta_0 + \beta_1 x + \ldots + \beta_{n_\beta} x^{n_\beta}, \qquad (4.70)$$

and the sequences of poles are the sequences of roots of the polynomial

$$\alpha_0 + \alpha_1 x + \ldots + \alpha_{n_\alpha} x^{n_\alpha}. \qquad (4.71)$$

The roots of (4.70), i.e. the zeros are denoted by $z_1(\tau), \ldots, z_{n_\beta}(\tau)$. The roots of (4.71), i.e. the poles are denoted by $p_1(\tau), \ldots, p_{n_\alpha}(\tau)$.

One of the results of Galois theory is that there is no formula for the roots of a fifth or higher degree polynomial equation in terms of the coefficients of the polynomial, using only the usual algebraic operations and application of radicals [Jacobson, 1985]. Mapping between the coefficient space and the space of roots is complicated, the roots behave very irregularly and indexing is a challenging task. Let us denote the continuous delay by $\tau_{\mathbb{R}}$. τ is nothing else than the sampled version of $\tau_{\mathbb{R}}$. $\theta_*(\tau_{\mathbb{R}})$ denotes a global optimum of $C(\theta, \tau_{\mathbb{R}})$. Although $\theta_*(\tau_{\mathbb{R}})$ is not a continuous function, it contains only a countable number of jump

discontinuities. Even if no multiple poles occur and $\theta(\tau_\mathbb{R})$ is a continuous function in an appropriate interval of $\mathbb{R}_+$, the roots of polynomials cannot be extracted as a function of $\tau_\mathbb{R}$ from the sampled roots, i.e. from $p_k(\tau)$, $k = 1, \ldots, n_\alpha$, $z_l(\tau)$, $l = 1, \ldots, n_\beta$. Hence, it is impossible to ensure that in the case $k \neq l$, $p_k(\tau)$ and $p_l(\tau)$ belong to the same $p_m(\tau_\mathbb{R})$ where $k, l, m = 1, \ldots, n_\alpha$.

To avoid such problems, only well selected subsequences are considered. We know from elementary analysis that every bounded sequence has a convergent subsequence. In the complex plane which can be mapped into $\mathbb{R}^3$ as a sphere (the Riemann sphere, [Rudin, 1987]), there is only one infinity and the sphere is closed and bounded (Heine-Borel theorem, [Rudin, 1991]), therefore it is true that every sequence contains a convergent subsequence. Let the set of positive integer numbers be denoted by $\mathbb{Z}_+$. The possible values of τ is the whole set, i.e. $\tau = \mathbb{Z}_+$. If $\tau' \subset \mathbb{Z}_+$ then $\theta(\tau')$ defines a subsequence of $\theta(\tau)$. Therefore, we use τ' to denote not only the subset, but $\theta(\tau')$ is used to denote the corresponding subsequence of $\theta(\tau)$. It is a little abuse of notation but will not cause problems. Similarly, convergent subsequences $z(\tau')$ and $p(\tau')$ can be constructed from the roots of the corresponding polynomials (4.70) and (4.71).

For a parameter vector sequence $\theta(\tau)$ can be constructed a disjoint composition of $\mathbb{Z}_+$

$$\mathbb{Z}_+ = \cup_{k=1}^{N_k} \tau_k \tag{4.72}$$

where N_k denotes the number of subsets and τ_k are the corresponding subset, such that

- for all $l = 1, \ldots, n_\beta$, $z_l(\tau_k)$ is convergent,

- for all $l = 1, \ldots, n_\alpha$, $p_l(\tau_k)$ is convergent,

- for all $l = 1, \ldots, n_\alpha$, if $\tau_{k,1}, \tau_{k,2} \in \tau_k$, $\tau_{k,1} \neq \tau_{k,2}$ then either $p_l(\tau_{k,1})$ and $p_l(\tau_{k,2})$ are stable, or $p_l(\tau_{k,1})$ and $p_l(\tau_{k,2})$ are unstable,

- for all $l = 1, \ldots, n_\beta$, if $\tau_{k,1}, \tau_{k,2} \in \tau_k$, $\tau_{k,1} \neq \tau_{k,2}$ then either $z_l(\tau_{k,1})$ and $z_l(\tau_{k,2})$ are inside the unit disk, or $z_l(\tau_{k,1})$ and $z_l(\tau_{k,2})$ are inside the complementary set of the unit disk.

It is worth noting that N_k can be infinity and it is possible for the cardinality of τ_k to be finite. If $|\tau_k|$ is finite, then by definition every pole sequence $p_k(\tau_k)$ and every zero sequence $z_l(\tau_l)$ is convergent. The last condition says that in a subsequence defined by τ_k the stability or the unstability of a pole cannot change.

In the rest of this chapter where the asymptotic value of a sequence of poles or zeros is referred, one of the convergent subsequence defined above is considered. Since most of the steps are examining the limits of the sequences or the limits of the function of the sequences, it will not cause problems.

In the following steps $\theta(\tau)$, $p_k(\tau)$ and $z_l(\tau)$ denote convergent sequences on the complex plane. Therefore, in asymptotic discussions, divergent sequences whose limit sets contain more than one point, will not occur in the rest of the chapter.

The functions, especially the transfer functions considered here are meromorphic on the complex plane. Hence, calculating the residue at a convergent subsequence of $p_1(\tau')$ leads

to a convergent sequence $\text{Res}_{z=p_1(\tau')}H(\theta(\tau'),z)$ where $H(\theta(\tau'),z)$ a meromorphic function having pole at $p_1(\tau')$.

The method used in the next subsection are based on the following. First, we assume that $\theta_*(\tau)$ is a parameter vector sequence which minimizes the cost function (4.1). Applying the introduced composition of τ, it is possible that for a τ_k, $\theta(\tau_k)$ defines an unstable transfer function. It means that it exists an integer $1 \leq l \leq n_\alpha$ such that $p_l(\tau_k)$ is unstable. In every case a new pole sequence $\tilde{p}_l(\tau_k)$ is constructed, which defines a new parameter vector sequence $\tilde{\theta}(\tau_k)$, such that

- for all τ_k, $\tilde{p}_l(\tau_k)$ is stable,

- for all τ_k sufficiently large $C(\tilde{\theta}(\tau_k),\tau_k) < C(\theta_*(\tau_k),\tau_k)$.

Hence, it is a contradiction: $\theta_*(\tau_k)$ cannot be the subsequence of the solution of the optimization problem. The ultimate conclusion will be that for all τ sufficiently large $\theta_*(\tau)$ defines stable systems.

4.4.4 Evaluating the Term $T_3(\theta)$

In this section extreme configurations of zeros and poles are examined. If the subset containing the poles and zeros which are converging to the complex infinity is not empty then contributions of the term $T_2(\theta,\tau)$ and $T_3(\theta)$ in the original cost function (4.43) are hard to analyst. This part of the proof is devoted to show that using $\overline{C}(\theta,\tau)$ it is possible to calculate the effects of the extreme subset. It will be proved that, asymptotically, the contributions of these poles and zeros to the numerator and the denominator of (4.69), are respectively the same. Hence, during the investigation of $\overline{C}(\theta,\tau)$ they can be left out.

First, a sufficient but not necessary condition is presented. We will prove that if $\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega$ converges to 0 then at least one pole converges to the (complex) infinity. Next, assuming that a group of poles and zeros are approaching infinity, the asymptotic value of $\overline{C}(\theta,\tau)$ is calculated.

4.4.4.1 Necessary Condition for $\| \widehat{H}(\Omega,\theta) \|_2$ Converging to 0

We assume that $\exists M$ such that $|p_k| < M$ for all $k = 1,\ldots,n_\alpha$. Substituting the polynomials into $\widehat{H}(e^{i\omega},\theta)$ and we get

$$\int_I \left|\widehat{H}(e^{i\omega},\theta)\right|^2 d\omega = \int_I \left|\frac{\prod_{k=1}^{n_\beta}(e^{i\omega}-z_k)}{\prod_{k=1}^{n_\alpha}(e^{i\omega}-p_k)}\right|^2 d\omega. \tag{4.73}$$

From the assumption we have $\omega \in I : \left|e^{i\omega}-p_k\right| \leq M+1$ and therefore we conclude that

$$\int_I \left|\widehat{H}(e^{i\omega},\theta)\right|^2 d\omega \geq \frac{1}{(M+1)^{2n_\alpha}} \int_I \left|\prod_{k=1}^{n_\beta}(e^{i\omega}-z_k)\right|^2 d\omega. \tag{4.74}$$

The interval $I = [0,2\pi]$ is divided into $2n_\beta + 1$ equal parts where the endpoints of the subintervals are denoted by I_l, $l = 1,\ldots,2n_\beta + 1$. Using the fact that the integral is

additive the last equation leads to

$$= \frac{1}{(M+1)^{2n_\alpha}} \sum_{l=1}^{2n_\beta+1} \int_{I_l} \left| \prod_{k=1}^{n_\beta} (e^{i\omega} - z_k) \right|^2 d\omega. \tag{4.75}$$

The pigeonhole principle ensures that there always exists an n such that none of the zeros converges to the arc defined as the image set of I_n of the function $e^{j\omega}$. Let us define the constant K_Z as

$$K_Z = \min \left\{ K_{Z,1}, \ldots, K_{Z,n_\beta} \right\} \tag{4.76}$$

where

$$K_{Z,m} = \arg \min_{\omega \in I_n} |e^{j\omega} - z_m|. \tag{4.77}$$

So, (4.75) can be continued

$$\geq \frac{1}{(M+1)^{2n_\alpha}} \int_{I_l} \left| \prod_{k=1}^{n_\beta} (e^{i\omega} - z_k) \right|^2 d\omega \geq \frac{1}{(M+1)^{2n_\alpha}} K_Z^2 \frac{2\pi}{2n_\beta + 1} \tag{4.78}$$

which is independent of τ. Therefore we conclude that if the poles remains finite then $\int_I \left| \widehat{H}(e^{j\omega}, \theta) \right|^2 d\omega$ does not converge to zero.

A simple construction shows that the condition that a pole converges to ∞ is not sufficient to have that $\int \left| \widehat{H}(e^{j\omega}, \theta) \right|^2 d\omega \to 0$. Let the orders of the polynomials be $n_\beta = 1$, $n_\alpha = 1$ and

$$z_1(\tau) = \tau + 1,$$
$$p_1(\tau) = \tau. \tag{4.79}$$

For all τ, the pole p_1 and the zero z_1 do not coincide, so the orders cannot be decreased. Substituting (4.79) into the expression of the integrand of $T_3(\theta)$ (4.64) leads to

$$\widehat{H}(z, \theta) \frac{1}{z} \widehat{H} \left(\frac{1}{z}, \theta \right) = \frac{1}{z} \frac{(z - z_1) \left(\frac{1}{z} - z_1 \right)}{(z - p_1) \left(\frac{1}{z} - p_1 \right)}$$

$$= \frac{z_1}{p_1} \frac{1}{z} \frac{(z - z_1) \left(\frac{1}{z_1} - z \right)}{(p_1 - z) \left(z - \frac{1}{p_1} \right)} \tag{4.80}$$

This complex function has two poles inside the unit circle. Therefore, following the same

lines of (4.52)

$$\frac{1}{2\pi} \int_I \left| \widehat{H}(e^{i\omega},\theta) \right|^2 d\omega = \mathrm{Res}_{z=0} \widehat{H}(z,\theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z},\theta\right) + \mathrm{Res}_{z=\frac{1}{p_1}} \widehat{H}(z,\theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z},\theta\right)$$

$$= \frac{z_1}{p_1} + \frac{z_1}{p_1} p_1 \frac{\left(\frac{1}{p_1} - z_1\right)\left(\frac{1}{z_1} - \frac{1}{p_1}\right)}{p_1 - \frac{1}{p_1}} = \frac{z_1}{p_1} + \frac{(1 - z_1 p_1)(p_1 - z_1)}{p_1(p_1^2 - 1)}$$

$$= \frac{\tau + 1}{\tau} + \frac{(1 - \tau(\tau + 1))(-1)}{\tau(\tau^2 - 1)} \to 1 \text{ as } \tau \to \infty. \tag{4.81}$$

4.4.4.2 Handling Poles and Zeros Converging to Infinity

Without any loss of generality we can index the poles and zeros as follows. The first L_D poles converge to infinity:

$$|p_l| \to \infty, \text{ if } l = 1, \ldots, L_D \text{ and}$$
$$\exists M_D < \infty \text{ such that } |p_l| < M_D \text{ for } l = L_D + 1, \ldots, n_\alpha$$

and the first L_N zeros converge to the infinity:

$$|z_l| \to \infty, \text{ if } l = 1, \ldots, L_N \text{ and}$$
$$\exists M_N < \infty \text{ such that } |z_l| < M_N \text{ for } l = L_N + 1, \ldots, n_\beta.$$

The bounded sequences of the poles are divided into two subsets:

$$|p_l| \to 0, \text{ if } l = L_D + 1, \ldots L_D + L_{D,0} \text{ and}$$
$$|p_l| \to K_{P,l} \text{ where } 0 < K_{P,l} < \infty \text{ for } l = L_D + L_{D,0} + 1, \ldots, n_\alpha$$

In order to simplify the notations we introduce the sets $\mathcal{P}_0$, $\mathcal{P}_\infty$ and $\mathcal{P}_D$ which contain only stable poles:

$$\mathcal{P}_\infty = \left\{ \frac{1}{p_1}, \ldots, \frac{1}{p_{L_D}} \right\}, \tag{4.82}$$

$$\mathcal{P}_0 = \left\{ p_{L_D+1} \ldots, p_{L_D+L_{D,0}} \right\}, \tag{4.83}$$

$$\mathcal{P}_D = \begin{cases} p_k & \text{if } p_k \text{ stable} \\ \dfrac{1}{p_k} & \text{if } p_k \text{ unstable} \end{cases} \quad k = L_D + L_{D,0} + 1, \ldots n_\alpha. \tag{4.84}$$

We define the following sequences: $\tilde{p}_l = \frac{1}{p_l}$ where $l = 1, \ldots, L_D$. By definition $\tilde{p}_l \to 0$. Using this definition, the term T_3/c_H^2 can be written as

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \sum_k \operatorname{Res}_{z=p_k, p_k \in \mathcal{P}_D} f(z)$$

$$+ \sum_k \operatorname{Res}_{z=p_k, p_k \in \mathcal{P}_0} f(z) + \sum_k \operatorname{Res}_{z=p_k, p_k \in \mathcal{P}_\infty} f(z) \tag{4.85}$$

where

$$f(z) = z^{n_\alpha - n_\beta - 1} \frac{\displaystyle\prod_{k=1}^{L_N} (z - z_k)(1 - z z_k) \prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k)}{\displaystyle\prod_{k=1}^{L_D} \left(z - \frac{1}{\tilde{p}_k}\right)(\tilde{p}_k - z) \prod_{k=L_D+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \prod_{k=1}^{L_D} \tilde{p}_k$$

$$= z^{n_\alpha - n_\beta - 1} \frac{\displaystyle\prod_{k=1}^{L_N} (z - z_k)(1 - z z_k) \prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k)}{\displaystyle\prod_{k=1}^{L_D} \left(\frac{1}{\tilde{p}_k} - z\right)(z - \tilde{p}_k) \prod_{k=L_D+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \prod_{k=1}^{L_D} \tilde{p}_k \tag{4.86}$$

and the sum extends over the set of poles which are inside the unit circle. $\widehat{H}_1(z, \theta)$ contains the poles and zeros of $\widehat{H}(z, \theta)$ that do not converge to infinity:

$$\widehat{H}_1(z, \theta) = \widehat{H}(z, \theta) \frac{\displaystyle\prod_{k=1}^{L_D} (z - p_k)}{\displaystyle\prod_{k=1}^{L_N} (z - z_k)} = \frac{\displaystyle\prod_{k=L_N+1}^{n_\beta} (z - z_k)}{\displaystyle\prod_{k=L_D+1}^{n_\alpha} (z - p_k)}. \tag{4.87}$$

The order of the numerator is $n_\beta - L_N$ and the order of the denominator is $n_\alpha - L_D$.

In the rest of this section we will prove that, asymptotically, in expression $\frac{1}{2\pi} \parallel \widehat{H}(\Omega, \theta) \parallel_2^2$ the contribution of the factor $\widehat{H}_1(z, \theta)$ and the contributions of the poles and the zeros converging to infinity can be handled independently as

$$\lim_{\tau \to \infty} \frac{\dfrac{1}{2\pi} \displaystyle\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega}{\displaystyle\prod_{k=1}^{L_N} z_k^2 \prod_{k=1}^{L_D} \tilde{p}_k^2 \frac{1}{2\pi} \int_I |\widehat{H}_1(\Omega, \theta)|^2 d\omega} = 1. \tag{4.88}$$

Using the previously proved necessary condition we know that $\int_I |\widehat{H}_1(\Omega, \theta)|^2 d\omega$ has a strictly positive lower bound and, therefore, the important terms in (4.85) are those which

do not have the property $o\left(\prod_{k=1}^{L_N} z_k^2 \prod_{k=1}^{L_D} \tilde{p}_k^2\right)$ as $\tau \to \infty$.

4.4.4.3 Contribution of the Residues of the Poles Converging to a Non-zero Value

The residue at the poles from the set $\mathcal{P}_D$ are studied, i.e. they are converging to a non-zero value inside the unit circle. We assume that $p_{l+L_D+L_{D,0}}$, $l = 1, \ldots, J$ where J is the number of poles which converge to the same point inside the unit circle. This set of poles is denoted by $\mathcal{P}_J$ and each may have a multiplicity larger than one. Now the contribution of the residues at poles from the $\mathcal{P}_J$ is studied. Using the mean value theorem for divided difference there exists a pole q_J with multiplicity J such that

$$\sum_{p_k \in \mathcal{P}_J} \text{Res}_{z=p_k} f(z) = \text{Res}_{z=q_J} f_J(z) \tag{4.89}$$

where

$$f_J(z) = \frac{z^{n_\alpha - n_\beta - 1} \prod_{k=1}^{L_N} (z - z_k)(1 - zz_k) \prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - zz_k)}{\prod_{k=1}^{L_D} \left(\frac{1}{\tilde{p}_k} - z\right)(z - \tilde{p}_k) \prod_{k=L_D+1}^{L_D+L_{D,0}} (z - p_k)(1 - p_k z)}$$

$$\times \frac{\prod_{k=1}^{L_D} \tilde{p}_k}{(z - q_J)^J \prod_{k=L_D+L_{D,0}+1}^{L_D+L_{D,0}+J} (1 - p_k z) \prod_{k=L_D+L_{D,0}+J+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \tag{4.90}$$

Applying the chain rule of differentiation to $f_J(z) = f_1(z)f_2(z)$ gives

$$[f_J(z)(z - q_J)^J]^{(J-1)} = [f_1(z)f_2(z)]^{(J-1)} = \sum_{k=0}^{J-1} \frac{(J-1)!}{(J-1-k)!k!} [f_1(z)]^{(k)} [f_2(z)]^{(J-1-k)}$$

$$= f_1(z)[f_2(z)]^{J-1} + \sum_{k=1}^{J-1} \frac{(J-1)!}{(J-1-k)!k!} [f_1(z)]^{(k)} [f_2(z)]^{(J-1-k)}. \tag{4.91}$$

It will be proved that the sum in (4.91) has asymptotically no contribution. Substituting

$$f_1(z) = z^{L_D - L_N} \frac{\prod_{k=1}^{L_N} (z - z_k)(1 - zz_k)}{\prod_{k=1}^{L_D} \left(\frac{1}{\tilde{p}_k} - z\right)(z - \tilde{p}_k)} \prod_{k=1}^{L_D} \tilde{p}_k = \frac{\prod_{k=1}^{L_N} (z - z_k)\left(\frac{1}{z} - z_k\right)}{\prod_{k=1}^{L_D} \left(\frac{1}{\tilde{p}_k} - z\right)\left(1 - \frac{\tilde{p}_k}{z}\right)} \prod_{k=1}^{L_D} \tilde{p}_k \quad (4.92)$$

and

$$f_2(z) = z^{n_\alpha - n_\beta - 1 - L_D + L_N} \frac{\prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - zz_k)}{\prod_{k,p_k \in (\mathcal{P}_0 \cup \mathcal{P}_D \setminus \mathcal{P}_J)} (z - p_k)(1 - p_k z) \prod_{k,p_k \in \mathcal{P}_J} (1 - p_k z)}. \quad (4.93)$$

Substituting $z = q_J$ into $f_1(z)$ we have the following factors

$$(q_J - z_l)\left(\frac{1}{q_J} - z_l\right) = z_l^2 + O(z_l) \quad l = 1, \ldots, L_N \quad (4.94)$$

and

$$\left(\frac{1}{\tilde{p}_l} - q_J\right)\left(1 - \frac{\tilde{p}_l}{q_J}\right) = \frac{1}{\tilde{p}_l} + O(1) \quad l = 1, \ldots, L_D. \quad (4.95)$$

Therefore, the overall asymptotic contribution of $f_1(q_J)$ is

$$\frac{f_1(z)|_{z=q_J}}{\prod_{k=1}^{L_N} z_k^2 \prod_{k=1}^{L_D} \tilde{p}_k^2} \to 1 \text{ as } \tau \to \infty. \quad (4.96)$$

Let us now study the factor $[f_2(z)]^{(J-1)}$ in (4.91). Using the facts that $|\mathcal{P}_0 \cup \mathcal{P}_D \setminus \mathcal{P}_J| = n_\alpha - L_D - J$, $|\mathcal{P}_J| = J$ and rewriting (4.93) we have

$$f_2(z) = \frac{1}{z} \frac{\prod_{k=L_N+1}^{n_\beta} (z - z_k)\left(\frac{1}{z} - z_k\right)}{\prod_{k,p_k \in (\mathcal{P}_0 \cup \mathcal{P}_D \setminus \mathcal{P}_J)} (z - p_k)\left(\frac{1}{z} - p_k\right) \prod_{k,p_k \in \mathcal{P}_J} \left(\frac{1}{z} - p_k\right)} \quad (4.97)$$

and applying the residue theorem gives

$$\frac{1}{(J-1)!} [f_2(z)]^{(J-1)}\Big|_{z=q_J} = \text{Res}_{z=q_J} \frac{f_2(z)}{(z - q_J)^J}. \quad (4.98)$$

If $[f_2(z)]^{(J-1)}\big|_{z=q_J} \to 0$ as $\tau \to \infty$ then it follows from (4.89), (4.91) and (4.96) that

$$\frac{\displaystyle\sum_{k,p_k\in\mathcal{P}_J} \mathrm{Res}_{z=p_k} f(z)}{\displaystyle\prod_{k=1}^{L_D} \tilde{p}_k^2 \prod_{k=1}^{L_n} z_k^2} \to 0, \tag{4.99}$$

so we do not need to examine this case deeper. Otherwise for all $p_k \in \mathcal{P}_J$

$$\left(\frac{1}{z} - p_k\right)\bigg|_{z=q_J} = \left(\frac{1}{z} - q_J\right)\bigg|_{z=q_J} + O\left(|q_J - p_k|\right). \tag{4.100}$$

Our assumption is that p_k converges a non-zero value and, hence, for all $p_k \in \mathcal{P}_J$, $\frac{1}{q_J} - p_k$ does not converge to zero and so does $\frac{1}{q_J} - q_J$. Therefore

$$[f_2(z)]^{(J-1)}\bigg|_{z=q_j} = \left[\frac{1}{z}\frac{\displaystyle\prod_{k=L_N+1}^{n_\beta}(z-z_k)\left(\frac{1}{z}-z_k\right)}{\displaystyle\prod_{k,p_k\in(\mathcal{P}_0\cup\mathcal{P}_D\setminus\mathcal{P}_J)}(z-p_k)\left(\frac{1}{z}-p_k\right)\left(\frac{1}{z}-q_J\right)^J}\right]^{(J-1)}\Bigg|_{z=q_j}$$

$$+O\left(\sum_{k,p_k\in\mathcal{P}_J}|q_J-p_k|\right). \tag{4.101}$$

We define $f_3(z)$ as

$$f_3(z) = \left[\frac{1}{z}\frac{\displaystyle\prod_{k=L_N+1}^{n_\beta}(z-z_k)\left(\frac{1}{z}-z_k\right)}{\displaystyle\prod_{k,p_k\in(\mathcal{P}_0\cup\mathcal{P}_D\setminus\mathcal{P}_J)}(z-p_k)\left(\frac{1}{z}-p_k\right)\left(\frac{1}{z}-q_J\right)^J}\right]^{(J-1)} \tag{4.102}$$

Hence, asymptotically we can write

$$\lim_{\tau\to\infty} \mathrm{Res}_{z=q_J}\frac{f_3(z)}{(z-q_J)^J} = \lim_{\tau\to\infty}\sum_{k,p_k\in\mathcal{P}_J} \mathrm{Res}_{z=p_k}\widehat{H}_1(z,\theta)\frac{1}{z}\widehat{H}_1\left(\frac{1}{z},\theta\right). \tag{4.103}$$

Next, the analysis of the sum in (4.91) is presented. Using the identity $\left(\frac{1}{\tilde{p}_k} - z\right) =$

$\frac{1}{\tilde{p}_k}(1 - z\tilde{p}_k)$, $f_1(z)$ can be written as

$$f_1(z) = \frac{\prod\limits_{k=1}^{L_N}(z - z_k)\left(\frac{1}{z} - z_k\right)}{\prod\limits_{k=1}^{L_D}(1 - \tilde{p}_k z)\left(1 - \frac{\tilde{p}_k}{z}\right)} \prod\limits_{k=1}^{L_D} \tilde{p}_k^2 \tag{4.104}$$

If $m > 0$ then

$$f_1(z)^{(m)} = \left[\frac{f_{1,N}(z)}{f_{1,D}(z)}\right]^{(m)} = \sum_{k=0}^{m} \frac{m!}{(m-k)!k!}\left(f_{1,N}(z)\right)^{(k)}\left(\frac{1}{f_{1,D}(z)}\right)^{(m-k)}. \tag{4.105}$$

In $f_{1,N}(z)$ differentiating the factors leads to

$$(z - z_l)^{(k)} = \begin{cases} 1, & \text{if } k = 1 \\ 0, & \text{if } k > 1 \end{cases} \tag{4.106}$$

and

$$\left(\frac{1}{z} - z_l\right)^{(k)} = \frac{1}{z^{k+1}}(-1)^k. \tag{4.107}$$

Our conclusion is

$$(z - z_l)^{(k)}\Big|_{z=q_J} = o(1) \text{ and } \left(\frac{1}{z} - z_l\right)^{(k)}\Big|_{z=q_J} = O(1). \tag{4.108}$$

Similarly,

$$\left[\frac{1}{(1 - z\tilde{p}_k)}\right]^{(m)} = \left[\frac{1}{(1 - z\tilde{p}_k)^2}\frac{\tilde{p}_k}{z^2}\right]^{(m-1)} = \frac{1}{(1 - z\tilde{p}_k)^{m+1}}\tilde{p}_k^m(m-1)! \tag{4.109}$$

and

$$\left[\frac{1}{\left(1 - \frac{\tilde{p}_k}{z}\right)}\right]^{(m)} = \frac{(-1)^{m-1}\tilde{p}_k}{(z - \tilde{p}_k)^{m+1}}(m-1)! \tag{4.110}$$

so

$$\left[\frac{1}{(1 - z\tilde{p}_k)}\right]^{(m)}\Big|_{z=q_J} = O\left(\tilde{p}_k^m\right) \tag{4.111}$$

$$\left[\frac{1}{\left(1-\frac{1}{\tilde{p}_k}\right)}\right]^{(m)}\Bigg|_{z=q_J} = O\left(\tilde{p}_k\right). \tag{4.112}$$

Therefore

$$m > 0 \Rightarrow f_1(z)^{(m)} = o\left(\prod_{k=1}^{L_N} z_k^2 \prod_{k=1}^{L_D} \tilde{p}_k^2\right). \tag{4.113}$$

In the end the arguments leads to

$$\lim_{\tau \to \infty} \frac{\displaystyle\sum_{k,p_k \in \mathcal{P}_J} \mathrm{Res}_{z=p_k} \widehat{H}(z,\theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z},\theta\right)}{\displaystyle\prod_{k=1}^{L_D} \tilde{p}_k^2 \prod_{k=1}^{L_N} z_k^2 \sum_{k,p_k \in \mathcal{P}_J} \mathrm{Res}_{z=p_k} \widehat{H}_1(z,\theta) \frac{1}{z} \widehat{H}_1\left(\frac{1}{z},\theta\right)} = 1 \tag{4.114}$$

4.4.4.4 Contributions of Residues calculated at Poles Converging to Zero

Now, the poles of (4.86) which are converging 0 or ∞ are examined. The number of poles converging to 0 depends also on the number $n_\alpha - n_\beta - 1$. (4.86) can be written as

$$= \frac{z^{n_\alpha - n_\beta - 1} \displaystyle\prod_{k=1}^{L_N} (z - z_k)\left(\frac{1}{z_k} - z\right) \prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k) \prod_{k=1}^{L_D} \tilde{p}_k \prod_{k=1}^{L_N} z_k}{\displaystyle\prod_{k,p_k \in \mathcal{P}_\infty} \left(\frac{1}{\tilde{p}_k} - z\right)(z - \tilde{p}_k) \prod_{k,p_k \in \mathcal{P}_0} (z - p_k)(1 - p_k z) \prod_{k,p_k \in \mathcal{P}_D} (z - p_k)(1 - p_k z)} \tag{4.115}$$

If $n_\alpha - n_\beta - 1 \geq 0$ then only the poles which are the inverse of $p_k \in \mathcal{P}_\infty$ and the poles form $\mathcal{P}_0$ are converging to 0. The cardinality of these sets are L_D and $L_{D,0}$, respectively. In the case where $n_\alpha - n_\beta - 1 < 0$ the number of poles in (4.115) which are converging to 0 are $n_\beta + 1 - n_\alpha + L_D + L_{D,0}$.

The following case are considered:

- $n_\alpha - n_\beta - 1 \geq 0$ and $n_\alpha - n_\beta - 1 \geq L_D + L_{D,0} - L_N$,

- $n_\alpha - n_\beta - 1 \geq 0$ and $n_\alpha - n_\beta - 1 - L_D - L_{D,0} + L_N < 0$,

- $n_\alpha - n_\beta - 1 < 0$.

- In the first case our assumption is $n_\alpha - n_\beta - 1 \geq L_D + L_{D,0} - L_N$. It means that only poles from the set $\mathcal{P}_0$ and inverse of poles from the set $\mathcal{P}_\infty$ converge to 0. The complex mean value theorem implies that there exists a q_0 for which

$$\sum_{k,p_k \in \mathcal{P}_0} \mathrm{Res}_{z=p_k} f(z) + \prod_{k,p_k \in \mathcal{P}_\infty} \mathrm{Res}_{z=\frac{1}{p_k}} f(z) = \mathrm{Res}_{z=q_0} f_0(z) \tag{4.116}$$

and its multiplicity is $L_D + L_{D,0} > 0$ where

$$f_0(z) = f(z) \frac{\prod\limits_{k, p_k \in \mathcal{P}_0} (z - p_k) \sum\limits_{k, p_k \in \mathcal{P}_\infty} \left(z - \frac{1}{p_k}\right)}{(z - q_0)^{L_D + L_{D,0}}}. \tag{4.117}$$

The function that has to be calculated at $z = q_0$ is

$$f_0(z)(z - q_0)^{L_D + L_{D,0}} = \frac{z^{n_\alpha - n_\beta - 1} \prod\limits_{k=1}^{L_N} (z - z_k) \left(\frac{1}{z_k} - z\right)}{\prod\limits_{k=1}^{L_D} \left(\frac{1}{\tilde{p}_k} - z\right) \prod\limits_{k=L_D+1}^{L_D + L_{D,0}+1} (1 - p_k z)}$$

$$\times \frac{\prod\limits_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k)}{\prod\limits_{k=L_D+L_{D,0}+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \prod\limits_{k=1}^{L_D} \tilde{p}_k \prod\limits_{k=1}^{L_N} z_k. \tag{4.118}$$

We define the following functions:

$$f_1(z) = z^{n_\alpha - n_\beta - 1} \tag{4.119}$$

and

$$f_2(z) = \frac{f(z)(z - q_0)^{L_D + L_{D,0}}}{f_1(z)}. \tag{4.120}$$

We known that

$$f_1(z)^{(m)} = \begin{cases} (n_\alpha - n_\beta - 1) \cdots (n_\alpha - n_\beta - 1 - m) z^{n_\alpha - n_\beta - 1 - m}, & \text{if } m < n_\alpha - n_\beta - 1, \\ (n_\alpha - n_\beta - 1)!, & \text{if } m = n_\alpha - n_\beta - 1, \\ 0 \text{ otherwise.} \end{cases}$$

Hence if $m \neq n_\alpha - n_\beta - 1$ then

$$q_0 \to 0 \ \Rightarrow \ f_1(z)^{(m)} \Big|_{z=q_0} \to 0 \text{ as } \tau \to \infty. \tag{4.121}$$

Since

$$\text{Res}_{z=q_0} f_0(z) = \frac{1}{(L_D + L_{D,0} - 1)!} \left[f_1(z) f_2(z)\right]^{(L_D + L_{D,0} - 1)} \Big|_{z=q_0}$$

$$= \frac{1}{(L_D + L_{D,0} - 1)!} \sum\limits_{k=0}^{L_D + L_{D,0} - 1} \frac{(L_D + L_{D,0} - 1)!}{(L_D + L_{D,0} - 1 - k)! k!}$$

$$\times f_1(z)^{(k)} f_2(z)^{(L_D+L_{D,0}-1-k)}\Big|_{z=q_0}$$

$$= \sum_{k=0}^{L_D+L_{D,0}-1} \frac{1}{(L_D+L_{D,0}-1-k)!k!} \left. f_1(z)^{(k)}\right|_{z=q_0} \left. f_2(z)^{(L_D+L_{D,0}-1-k)}\right|_{z=q_0} \quad (4.122)$$

Combining the assumption $n_\alpha - n_\beta - 1 \geq L_D + L_{D,0} - L_N$, (4.121), and (4.106) to (4.113), it can be seen that

$$\mathrm{Res}_{z=q_0} f_0(z) = o\left(\prod_{k=1}^{L_N} z_k^2 \prod_{k=1}^{L_D} \tilde{p}_k^2\right). \quad (4.123)$$

- The assumptions are $n_\alpha - n_\beta - 1 \geq 0$ and $n_\alpha - n_\beta - 1 - L_D - L_{D,0} + L_N < 0$. Like in the previous case the mean value theorem for divided difference is used and f_0 is defined as

$$\sum_{k,p_k \in \mathcal{P}_0} \mathrm{Res}_{z=p_k} f(z) + \sum_{k,p_k \in \mathcal{P}_\infty} \mathrm{Res}_{z=\frac{1}{p_k}} f(z) = \mathrm{Res}_{z=q_0} f_0(z) \quad (4.124)$$

$$f_0(z) = f(z)\frac{\displaystyle\prod_{k,p_k \in \mathcal{P}_0}(z-p_k) \sum_{k,p_k \in \mathcal{P}_\infty}\left(z-\frac{1}{p_k}\right)}{(z-q_0)^{L_D+L_{D,0}}}. \quad (4.125)$$

Since

$$\frac{1}{\left(\frac{1}{\tilde{p}_k}-z\right)} = \frac{\tilde{p}_k}{(1-z\tilde{p}_k)} \quad (4.126)$$

and

$$(z-z_k) = z_k\left(\frac{z}{z_k}-1\right) \quad (4.127)$$

we can write

$$\mathrm{Res}_{z=q_0} f_0(z) = \frac{1}{(L_D+L_{D,0}-1)!} \left[(z-q_0)^{L_D+L_{D,0}} f_0(z)\right]^{(L_D+L_{D,0}-1)}\Big|_{z=q_0}$$

$$= \frac{1}{(L_D+L_{D,0}-1)!} \left[\frac{z^{n_\alpha-n_\beta-1}\displaystyle\prod_{k=1}^{L_N}\left(1-\frac{z}{z_k}\right)\left(z-\frac{1}{z_k}\right)}{\displaystyle\prod_{k=1}^{L_D}(1-\tilde{p}_k z)}\right.$$

$$\times \left. \left[\frac{\displaystyle\prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - zz_k) \prod_{k=1}^{L_D} \tilde{p}_k^2 \prod_{k=1}^{L_N} z_k^2}{\displaystyle\prod_{k=L_D+1}^{L_D+L_{D,0}} (1 - p_k z) \prod_{k=L_D+L_{D,0}+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \right]^{(L_D+L_{D,0}-1)} \right|_{z=q_0} \quad (4.128)$$

We define the following functions:

$$f_1(z) = z^{n_\alpha - n_\beta - 1} \prod_{k=1}^{L_N} \left(z - \frac{1}{z_k} \right), \qquad (4.129)$$

$$f_2(z) = \frac{\displaystyle\prod_{k=1}^{L_N} \left(1 - \frac{z}{z_k} \right)}{\displaystyle\prod_{k=1}^{L_D} (1 - \tilde{p}_k z) \prod_{k=L_D+1}^{L_D+L_{D,0}} (1 - p_k z)} \qquad (4.130)$$

and

$$f_3(z) = \frac{\displaystyle\prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - zz_k)}{\displaystyle\prod_{k=L_D+L_{D,0}+1}^{n_\alpha} (z - p_k)(1 - p_k z)}. \qquad (4.131)$$

$$\frac{\mathrm{Res}_{z=q_0} f_0(z)}{\displaystyle\prod_{k=1}^{L_D} \tilde{p}_k^2 \prod_{k=1}^{L_N} z_k^2} = \frac{1}{(L_D + L_{D,0} - 1)!} \left[f_1(z) f_2(z) f_3(z) \right]^{(L_D+L_{D,0}-1)} \Big|_{z=q_0}. \quad (4.132)$$

Using the same arguments as in the previous case

$$f_1(z)^{(m)} \Big|_{z=q_0} = \begin{cases} o(1), & \text{if } m < n_\alpha - n_\beta - 1 + L_N, \\ (n_\alpha - n_\beta - 1)!, & \text{if } m = n_\alpha - n_\beta - 1 + L_N, \\ 0, & \text{otherwise.} \end{cases}$$

Now, the factors of $f_2(z)$ are studied. If $m > 0$ then

$$\left[\left(1 - \frac{z}{z_k} \right) \right]^{(m)} = \left\{ O\left(\frac{1}{z_k} \right), \text{ if } m = 1, \quad 0, \text{ if } m > 1, \right. \qquad (4.133)$$

where $|z_k| \to \infty$ as $\tau \to \infty$, and if $m = 0$

$$\left(1 - \frac{q_0}{z_k}\right) \to 1 (\tau \to \infty). \tag{4.134}$$

Contributions of the factors which belong to the set $\mathcal{P}_\infty$ are

$$\left[\frac{1}{(1 - \tilde{p}_k z)}\right]^{(m)}\Bigg|_{z=q_0} = O(\tilde{p}_k^m) \text{ for } k = 1, \ldots, L_D \tag{4.135}$$

where $\tilde{p}_k \to 0$ as $\tau \to \infty$, and contributions of the factors which belong to the set $\mathcal{P}_0$

$$\left[\frac{1}{(1 - p_k z)}\right]^{(m)} = O(p_k^m) \text{ for } k = L_D + 1, \ldots, L_D + L_{D,0} \tag{4.136}$$

where $p_k \to 0$ as $\tau \to \infty$. If $m = 0$ then

$$\lim_{\tau \to \infty} \left(\frac{1}{1 - \tilde{p}_k q_0}\right) = \lim_{\tau \to \infty} \left(\frac{1}{1 - p_k q_0}\right) = 1. \tag{4.137}$$

So, we get

$$f_2(z)^{(m)}\Big|_{z=q_0} = \begin{cases} o(1), & \text{if } m > 0, \\ 1, & \text{otherwise.} \end{cases} \tag{4.138}$$

Therefore, in (4.132)

$$= \frac{1}{(L_D + L_{D,0} - 1)!} \left(\frac{(L_D + L_{D,0} - 1)!}{(n_\alpha - n_\beta - 1 + L_N)!(L_D + L_{D,0} - n_\alpha + n_\beta + 1 - L_N)!}\right.$$

$$\times f_1(z)^{(n_\alpha - n_\beta - 1 + L_N)} [f_2(z) f_3(z)]^{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)}$$

$$+ \sum_{k=0, k \neq n_\alpha - n_\beta - 1 + L_N}^{L_D + L_{D,0} - 1} \frac{(L_D + L_{D,0} - 1)!}{(L_D + L_{D,0} - 1 - k)! k!}$$

$$\left.\times f_1(z)^{(k)} [f_2(z) f_3(z)]^{(L_D + L_{D,0} - 1 - k)}\right) \tag{4.139}$$

the first term has asymptotically importance. Moreover, from (4.138) we conclude that in the expression

$$[f_2(z) f_3(z)]^{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)} = f_2(z) f_3(z)^{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)}$$

$$+ \sum_{k=1}^{(L_D+L_{D,0}-n_\alpha+n_\beta-L_N)} \frac{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)!}{((L_D + L_{D,0} - n_\alpha + n_\beta - L_N) - k)!k!}$$

$$\times f_2(z)^{(k)} f_3(z)^{(L_D+L_{D,0}-n_\alpha+n_\beta-L_N-k)} \tag{4.140}$$

only the first term may converge to a non-zero value. The ultimate result of this case that, asymptotically, only the following term counts

$$f_1(z)^{(n_\alpha-n_\beta-1+L_N)} f_2(z) \frac{1}{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)!}$$

$$\times f_3(z)^{(L_D+L_{D,0}-n_\alpha+n_\beta-L_N)}. \tag{4.141}$$

$$\frac{1}{(L_D + L_{D,0} - n_\alpha + n_\beta - L_N)!}$$

$$\times \left[\frac{\prod_{k=L_N+1}^{n_\beta}(z - z_k)(1 - zz_k)}{\prod_{k=L_D+L_{D,0}+1}^{n_\alpha} (z - p_k)(1 - p_k z)} \right]^{(n_\beta-n_\alpha-L_D-L_{D,0}+L_N)} \Bigg|_{z=q_0}$$

$$= \operatorname{Res}_{z=q_0} \frac{\prod_{k=L_N+1}^{n_\beta}(z - z_k)\left(\frac{1}{z} - z_k\right)}{z\prod_{k=L_D+1}^{n_\alpha}(z - p_k)\left(\frac{1}{z} - p_k\right)} = \operatorname{Res}_{z=q_0} \widehat{H}_1(z,\theta)\frac{1}{z}\widehat{H}_1\left(\frac{1}{z},\theta\right) \tag{4.142}$$

Hence

$$\lim_{\tau \to \infty} \left(\sum_{k,p_k \in \mathcal{P}_0} \operatorname{Res}_{z=p_k} \widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) + \sum_{k,p_k \in \mathcal{P}_\infty} \operatorname{Res}_{z=\frac{1}{p_k}} \widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) \right)$$

$$\times \frac{1}{\prod_{k=1}^{L_D} \tilde{p}_k^2 \prod_{l=1}^{L_N} z_l^2 \operatorname{Res}_{z=q_0} \widehat{H}_1(z,\theta)\frac{1}{z}\widehat{H}_1\left(\frac{1}{z},\theta\right)} = 1 \tag{4.143}$$

- The case $n_\alpha - n_\beta - 1 < 0$. It follows from the assumptions that the factor $z^{n_\alpha-n_\beta-1} = \frac{1}{z^{n_\beta+1-n_\alpha}}$ has $n_\beta + 1 - n_\alpha$ number of poles at 0, therefore the number of the poles converging to 0 equals $n_\beta - n_\alpha + 1 + L_{D,0} + L_D$. The divided difference theorem says that there exists a sequence of $q_0(\tau)$ such that

$$\sum_{k,p_k \in \mathcal{P}_0} \operatorname{Res}_{z=p_k} f(z) + \sum_{k,p_k \in \mathcal{P}_\infty} \operatorname{Res}_{z=\frac{1}{p_k}} f(z) + \operatorname{Res}_{z=0} f(z) = \operatorname{Res}_{z=q_0} f_0(z) \tag{4.144}$$

where the multiplicity of q_0 is $n_\beta + 1 - n_\alpha + L_{D,0} + L_D$ and $|q_0(\tau)| \to 0$ as $\tau \to \infty$ and

$$f_0(z) = f(z) \frac{\displaystyle\prod_{k,p_k \in \mathcal{P}_0} (z - p_k) \sum_{k,p_k \in \mathcal{P}_\infty} \left(z - \frac{1}{p_k}\right) z^{n_\beta - n_\alpha + 1}}{(z - q_0)^{N+1}}.$$

$$\mathrm{Res}_{z=q_0} f_0(z) = \frac{1}{N!} \left[\frac{\displaystyle\prod_{k=1}^{L_N} \left(1 - \frac{z}{z_k}\right)\left(z - \frac{1}{z_k}\right)}{\displaystyle\prod_{j=1}^{L_D} (1 - \tilde{p}_j z) \prod_{k=L_D+1}^{L_D+L_{D,0}} (1 - p_k z)} \right.$$

$$\left. \times \frac{\displaystyle\prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k)}{\displaystyle\prod_{k=L_D+1+L_{D,0}}^{n_\alpha} (z - p_k)(1 - p_k z)} \prod_{j=1}^{L_D} \tilde{p}_j^2 \prod_{j=1}^{L_N} z_j^2 \right]^{(N)} \Bigg|_{z=q_0} \tag{4.145}$$

where $N = L_D + L_{D,0} - n_\alpha + n_\beta$. We define the functions $f_1(z)$, $f_2(z)$, $f_3(z)$ and $f_4(z)$ as

$$f_1(z) = \frac{\prod_{k=1}^{L_N} \left(1 - \frac{z}{z_k}\right)}{\prod_{k=1}^{L_D} (1 - \tilde{p}_k z)}, \tag{4.146}$$

$$f_2(z) = \prod_{k=1}^{L_N} \left(z - \frac{1}{z_k}\right), \tag{4.147}$$

$$f_3(z) = \frac{\prod_{k=L_N+1}^{n_\beta} (z - z_k)(1 - z z_k)}{\prod_{k=L_D+1+L_{D,0}}^{n_\alpha} (z - p_k)(1 - p_k z)}, \tag{4.148}$$

$$f_4(z) = \frac{1}{\prod_{k=L_D+1}^{L_D+L_{D,0}} (1 - p_k z)}. \tag{4.149}$$

The chain rule of differentiation leads to

$$f(z)^{(N)} = [f_1(z) f_2(z) f_3(z) f_4(z)]^{(N)} =$$

$$= (f_2(z)f_3(z))^{(N)}f_1(z)f_4(z) + \sum_{k=1}^{N} \frac{N!}{(N-k)!k!}(f_2(z)f_3(z))^{(N-k)}(f_1(z)f_4(z))^{(k)}$$

$$= \left(\frac{N!}{(N-L_N)!L_N!}f_2(z)^{(L_N)}f_3(z)^{(N-L_N)} \right.$$

$$+ \sum_{l=0,l\neq L_N}^{N} \frac{N!}{(N-l)!l!}f_2(z)^{(l)}f_3(z)^{(N-l)} \right) f_1(z)f_4(z)$$

$$+ \sum_{k=1}^{N} \frac{N!}{(N-k)!k!}(f_2(z)f_3(z))^{(N-k)}(f_1(z)f_4(z))^{(k)}. \tag{4.150}$$

If $m > 0$ is an integer then using the facts that $\left[\frac{1}{1-p_k z} \right]' = p_k \frac{1}{(1-p_k z)^2}$ and $p_k \to 0$ for all $p_k \in \mathcal{P}_0$ we have

$$\left. f_4(z)^{(m)} \right|_{z=q_0} \to 0 \text{ as } \tau \to \infty. \tag{4.151}$$

Otherwise because for all $p_k \in \mathcal{P}_0$ $\frac{1}{1-p_k q_0} \to 1$ as $\tau \to \infty$

$$\left. f_4(z) \right|_{z=q_0} \to 1 \text{ as } \tau \to \infty. \tag{4.152}$$

Similarly, if $m > 0$ then using the facts that $\left[1 - \frac{z}{z_k} \right]' = -\frac{1}{z_k}$, $z_k \to \infty$ for $k = 1, \ldots, L_N$ and $\left[\frac{1}{1-\tilde{p}_k z} \right]' = \tilde{p}_k \frac{1}{(1-\tilde{p}_k z)^2}$, $\tilde{p}_k \to 0$ for all $k : p_k \in \mathcal{P}_\infty$.

$$\left. f_1(z)^{(m)} \right|_{z=q_0} \to 0 \text{ as } \tau \to \infty. \tag{4.153}$$

and in case where $m = 0$

$$\left. f_1(z) \right|_{z=q_0} \to 1 \text{ as } \tau \to \infty. \tag{4.154}$$

If $m < L_N$ then $f_2(z)^{(m)}$ contains at least one of the factors of (4.147). Since $q_0 \to 0$ and $\frac{1}{z_k} \to 0$ for all $k = 1, \ldots, L_N$, so $\left(q_0 - \frac{1}{z_k} \right) \to 0$ therefore

$$m < L_N \quad \Rightarrow \quad \left. f_2(z)^{(m)} \right|_{z=q_0} \to 0 \text{ as } \tau \to \infty. \tag{4.155}$$

We know that

$$f_2(z)^{(L_N)} = L_N!, \tag{4.156}$$

and if $m > L_N$ then

$$f_2(z)^{(L_N)} = 0. \tag{4.157}$$

Our conclusion is that $f_2(z)^{(L_N)} f_3(z)^{(N-L_N)} f_1(z) f_4(z)\big|_{z=q_0}$ is the only term which may converge to non-zero as $\tau \to \infty$. The limit is non zero in the case where $H(z)$ does not contain any zero converging to 0. Since

$$f_2(z)^{(L_N)} = L_N!, \ f_1(q_0) \to 1, \ f_4(q_0) \to 1 \tag{4.158}$$

the limit of the first term in (4.150) can be calculated as

$$\frac{1}{(N-L_N)!} f_3(z)^{(N-L_N)} \bigg|_{z=q_0} = \frac{1}{(N-L_N)!} f_3(z)^{(L_D-L_N+L_{D,0}-n_\alpha+n_\beta)} \bigg|_{z=q_0}$$

$$= \frac{1}{(N-L_N)!} \left[\frac{\prod_{k=L_N+1}^{n_\beta} (z-z_k)(1-zz_k)}{\prod_{k=L_D+1+L_{D,0}}^{n_\alpha} (z-p_k)(1-p_kz)} \right]^{(L_D-L_N+L_{D,0}-n_\alpha+n_\beta)} \bigg|_{z=q_0}$$

$$= \mathrm{Res}_{z=q_0} \ z^{-(N-L_N+1)} \frac{\prod_{k=L_N+1}^{n_\beta} (z-z_k)(1-z_kz)}{\prod_{k=L_D+1+L_{D,0}}^{n_\alpha} (z-p_k)(1-p_kz)}$$

$$= \mathrm{Res}_{z=q_0} \ \frac{1}{z} \frac{\prod_{k=L_N+1}^{n_\beta} (z-z_k)\left(\frac{1}{z}-z_k\right)}{\prod_{j=L_D+1+L_{D,0}}^{n_\alpha} (z-p_j)\left(\frac{1}{z}-p_j\right)} \tag{4.159}$$

It means that asymptotically the following equality holds

$$\lim_{\tau \to \infty} \frac{1}{\frac{\prod_{k=1}^{L_N} z_k^2}{\prod_{k=1}^{L_D} p_k^2} \mathrm{Res}_{z=q_0} \frac{1}{z} \widehat{H}_1(z,\theta) \widehat{H}_1\left(\frac{1}{z},\theta\right)}$$

$$\times \left(\sum_{k,p_k \in \mathcal{P}_0} \mathrm{Res}_{z=p_k} \frac{1}{z} \widehat{H}(z,\theta) \widehat{H}\left(\frac{1}{z},\theta\right) + \sum_{k,p_k \in \mathcal{P}_\infty} \mathrm{Res}_{z=\frac{1}{p_k}} \frac{1}{z} \widehat{H}(z,\theta) \widehat{H}\left(\frac{1}{z},\theta\right) \right.$$

$$\left. + \mathrm{Res}_{z=0} \frac{1}{z} \widehat{H}(z,\theta) \widehat{H}\left(\frac{1}{z},\theta\right) \right) = 1. \tag{4.160}$$

4.4.4.5   Contribution to $T_2(\theta, \tau)$

In the term $T_2(\theta, \tau)$ the residues calculated at points inside the unit circle, see (4.51). Therefore the inverses of the poles from $\mathcal{P}_\infty$ play no role in the term $T_2(\theta, \tau)$. In the

numerator of $\overline{C}(\theta, \tau)$ in (4.69) the residues at the poles from $\mathcal{P}_D$, $\mathcal{P}_0$ are evaluated.

Let us assume that the poles of $\widehat{H}(z, \theta)$ and $T\left(\frac{1}{z}\right)$ which constitute a set denoted by $\mathcal{P}_q$ converge to the same point inside unit circle. Applying the divided difference theorem there exists a pole q with multiplicity J such that

$$\sum_{k, p_k \in \mathcal{P}_q} \text{Res}_{z=p_k} \widehat{H}(z, \theta) z^{\tau-1} T\left(\frac{1}{z}\right) = \text{Res}_{z=q} f_0(z) \tag{4.161}$$

where

$$f_0(z) = \widehat{H}(z, \theta) z^{\tau-1} T\left(\frac{1}{z}\right) \frac{\prod_{k, p_k \in \mathcal{P}_q} (z - p_k)}{(z - q)^J}. \tag{4.162}$$

Substituting $z = q$ it leads to

$$\text{Res}_{z=q} f_0(z) = \frac{1}{(J-1)!} \times$$

$$= \times \left[\frac{\prod_{k=1}^{L_N} (z - z_k) \prod_{k=L_N+1}^{n_\beta} (z - z_k) z^{\tau-1} \prod_{k=1}^{n_{\beta_T}} \left(\frac{1}{z} - z_{k,T}\right)}{\prod_{k=1}^{L_D} (z - p_k) \prod_{k, p_k \notin (\mathcal{P}_q \cup \mathcal{P}_\infty)} (z - p_k) \prod_{k, p_{T,k} \notin \mathcal{P}_q} \left(\frac{1}{z} - p_{T,k}\right)}\right]^{(J-1)} \Bigg|_{z=q}. \tag{4.163}$$

We define the following functions:

$$f_1(z) = \frac{\prod_{k=1}^{L_N}(z - z_k)}{\prod_{k=1}^{L_D}(z - p_k)} \tag{4.164}$$

and

$$f_2(z) = \frac{1}{(J-1)!} \frac{\prod_{k=1}^{L_N}(z - z_k) \prod_{k=L_N+1}^{n_\beta} (z - z_k) z^{\tau-1} \prod_{k=1}^{n_{\beta_T}} \left(\frac{1}{z} - z_{k,T}\right)}{\prod_{k, p_k \notin (\mathcal{P}_q \cup \mathcal{P}_\infty)} (z - p_k) \prod_{k, p_{T,k} \notin \mathcal{P}_q} \left(\frac{1}{z} - p_{T,k}\right)}. \tag{4.165}$$

The chain rule of differentiation gives

$$[f_1(z) f_2(z)]^{(J-1)} = f_1(z) f_2(z)^{(J-1)} + \sum_{k=1}^{J-1} \frac{(J-1)!}{(J-1-k)! k!} f_1(z)^{(k)} f_2(z)^{(J-1-k)}. \tag{4.166}$$

If $m > 1$ $f_1(q)^{(m)}$ contains less factors in the numerator or more factors in the denominator. Since $z_k \to \infty$ and $p_k \to \infty$ it means that the convergence rate is decreased. $|z_k| \to \infty$

$(\tau \to \infty)$, $|q| < 1$ we have

$$\lim_{\tau \to \infty} \left| \frac{\left[\prod_{k=1}^{L_N} (z - z_k) \right]^{(m_1)} \Big|_{z=q}}{\prod_{k=1}^{L_N} (q - z_k)} \right| = \begin{cases} 0, \text{ if } m_1 > 0, \\ 1, \text{ if } m_1 = 0, \end{cases}$$

and because of the fact that $|p_k| \to \infty$ we have

$$\lim_{\tau \to \infty} \left| \frac{\left[\prod_{k=1}^{L_N} \frac{1}{(z - p_k)} \right]^{(m_2)} \Big|_{z=q}}{\prod_{k=1}^{L_N} \frac{1}{(q - z_k)}} \right| = \begin{cases} 0, \text{ if } m_2 > 0, \\ 1, \text{ if } m_2 = 0. \end{cases}$$

Therefore asymptotically the term $f_1(z)f_2(z)^{(J-1)}$ is of importance. So

$$\frac{\prod_{k=1}^{L_N} (q - z_k)}{\prod_{k=1}^{L_D} (q - p_k)} = \frac{\prod_{k=1}^{L_N} z_k}{\prod_{k=1}^{L_D} p_k} (-1)^{L_N + L_D} \left(\frac{\prod_{k=1}^{L_N} \left(1 - \frac{q}{z_k} \right)}{\prod_{k=1}^{L_D} \left(1 - \frac{q}{p_k} \right)} \right)$$

$$= \frac{\prod_{k=1}^{L_N} z_k}{\prod_{k=1}^{L_D} p_k} (-1)^{L_N + L_D} \left(1 + O\left(\sum_{k=1}^{L_N} \frac{1}{z_k} + \sum_{k=1}^{L_D} \frac{1}{p_k} \right) \right)$$

$$= \frac{\prod_{k=1}^{L_N} z_k}{\prod_{k=1}^{L_D} p_k} (-1)^{L_N + L_D} (1 + o(1)). \tag{4.167}$$

Furthermore,

$$\lim_{\tau \to \infty} \frac{f_2(z)^{(J-1)} \Big|_{z=q}}{\sum_{k, p_k \in \mathcal{P}_q} \operatorname{Res}_{z=p_k} \widehat{H}_1(z, \theta) z^{\tau-1} T\left(\frac{1}{z} \right)} = 1 \tag{4.168}$$

and therefore

$$\lim_{\tau \to \infty} \frac{\sum_{k, p_k \in \mathcal{P}_q} \operatorname{Res}_{z=p_k} \widehat{H}(z, \theta) z^{\tau-1} T\left(\frac{1}{\tau} \right)}{\frac{\prod_{k=1}^{L_N} z_k}{(-1)^{L_N + L_D} \prod_{k=1}^{L_D} p_k} \sum_{k, p_k \in \mathcal{P}_q} \operatorname{Res}_{z=p_k} \widehat{H}_1(z, \theta) z^{\tau-1} T\left(\frac{1}{z} \right)} = 1 \tag{4.169}$$

where $\widehat{H}_1(z,\theta)$ contains only the poles and zeros of $\widehat{H}(z,\theta)$ that do not converge to the infinity.

4.4.4.6 Conclusion for $\overline{C}(\theta,\tau)$

Collecting the results of the previous subsections the term $T_2(\theta,\tau)$ can be written as

$$\left(\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}(\Omega,\theta)d\omega\right)^2 = \left(\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}_1(\Omega,\theta)d\omega\right)^2 \left(\frac{\prod_{k=1}^{L_N} z_k}{\prod_{k=1}^{L_D} p_k}(1+o(1))\right)^2$$

and using the usual rules with respecting to the function $o(.)$ (see 4.2.2) we continue

$$= \left(\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}_1(\Omega,\theta)d\omega\right)^2 \left(\frac{\prod_{k=1}^{L_N} z_k^2}{\prod_{k=1}^{L_D} p_k^2}(1+o(1))\right) \tag{4.170}$$

and the term $T_3(\theta)$ can be expressed as a sequence of $\theta(\tau)$:

$$\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \int_I |\widehat{H}_1(\Omega,\theta)|^2 d\omega \left(\frac{\prod_{k=1}^{L_N} z_k^2}{\prod_{k=1}^{L_D} p_k^2}(1+o(1))\right). \tag{4.171}$$

Therefore

$$\overline{C}(\theta,\tau) = \frac{\left(\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}_1(\Omega,\theta)d\omega\right)^2}{\int_I |\widehat{H}_1(\Omega,\theta)|^2 d\omega}(1+o(1)) \tag{4.172}$$

The last expression shows that if we consider the parameter $\theta_*(\tau)$ vector that determines the global minimum of the cost function then the zeros and the poles which are converging to ∞ play no role as $\tau \to \infty$. Therefore, from now on the transfer function is assumed to have no poles nor zeros converging to infinity.

As a consequence of §4.4.1 every term in (4.172) is bounded. In the sequel we will prove that in all cases $\overline{C}(\theta_*,\tau)$ converges to 0.

4.4.5 Existence of a Pole Converging to the Unit Circle

Now we prove that there exists at least one stable pole series whose absolute value converges to 1. We prove it again by contradiction. If it is not true then every pole inside the unit circle has a common bound K_2.

$$\forall \tau, \forall 0 \le k \le n_\alpha : |p_k| < K_2 < 1 \tag{4.173}$$

Because of the factor z^τ, $T_2(\theta, \tau)$ in (4.51) converges towards zero at least exponentially for $\tau \to \infty$. The numerator of (4.69) can be estimated as

$$\left(2\pi \sum_k \text{Res}_{z=p_k} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) \right)^2 \leq C_1 K_2^{2\tau} \tag{4.174}$$

where C_1 is an appropriate constant. We apply the result of the previous section. Using the fact that the denominator of (4.69) is bounded below and applying the estimation (4.174) we get

$$\overline{C}(\theta, \tau) \leq C_1 K_2^{2\tau} K_{T_3} \tag{4.175}$$

where K_{T_3} is the upper bound of the reciprocal of $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$. (4.68) and the last equation shows that the whole cost function converges towards zero at least exponentially.

Let $\sigma_{T,W}$ denote the number of zeros of $T(z)$ which are equal to w, and let $\sigma_{T_{-1},W}$ denote the number of zeros of $T\left(\frac{1}{z}\right)$ which are equal to w. Let us define $\sigma = \min\{\sigma_{T,1}, \sigma_{T,-1}\}$. We know that $\sigma_{T,1} = \sigma_{T_{-1},1}$ and $\sigma_{T,-1} = \sigma_{T_{-1},-1}$.

Now via another construction we are going to show that the function (4.69) converges towards zero at the rate $1/\tau^{2\sigma+1}$. Let us select a pole p_1 (or of course we can select a complex conjugate pair). Let the amplitude of this (these) be such that

$$p_1 = P_1 \left(1 - \frac{1}{\tau} \right) \tag{4.176}$$

where $P_1 = 1$ if $\sigma = \sigma_{T_{-1},1}$ and $P_1 = -1$ otherwise, and by construction there is no sub-sequence of zeros z_1 such that

$$|p_1 - z_1| = o(1). \tag{4.177}$$

For a a complex conjugate pole pair the angle of P is neither zero nor π.

$T\left(\frac{1}{z}\right)$ can be written as

$$T\left(\frac{1}{z}\right) = (z-1)^{\sigma_{T_{-1},1}} (z+1)^{\sigma_{T_1,-1}} \tilde{T}(z) \tag{4.178}$$

where $\tilde{T}(z)$ is the part of the target function for which $\tilde{T}(1) \neq 0$ and $\tilde{T}(-1) \neq 0$. $T_2(\theta, \tau)/c_H = O\left(\frac{1}{\tau^\sigma}\right)$ and by construction for any sequence $l_\tau \to \mathbb{T}$ the sequence $\tilde{T}^{-1}(l_\tau)$ is bounded as $\tau \to \infty$ because of note 4.2.7. Let us denote the maximum of the possible non-zero limit value $K_{\tilde{T}^{-1}}$.

The third term $T_3(\theta)$ (4.86) of the cost function contains a multiplicator c_H^2. If $c_H \neq 0$ then

$$\forall \theta : \frac{T_3(\theta(\tau))}{c_H^2} > 0. \tag{4.179}$$

According to result of the previous section $T_3(\theta)/c_H^2$ has a positive lower bound which is independent of τ.

Let us write the corresponding residue(s) of $T_3(\theta)$ at p_1 in the case of a real pole as

$$\operatorname{Res}_{z=p_1} H(z^{-1},\theta) z^{-1} H(z,\theta) = \operatorname{Res}_{z=p_1} \frac{1}{(z-p_1)\left(\dfrac{1}{z}-p_1\right)} \frac{1}{z} H_1(z,\theta) \tag{4.180}$$

or in the case of conjugate pair of poles as

$$\operatorname{Res}_{z=p_1} H(z^{-1},\theta) z^{-1} H(z,\theta) + \operatorname{Res}_{z=p_1^*} H(z^{-1},\theta) z^{-1} H(z,\theta)$$

$$= \operatorname{Res}_{z=p_1} \frac{1}{(z-p_1)\left(\dfrac{1}{z}-p_1\right)(z-p_1^*)\left(\dfrac{1}{z}-p_1^*\right)} \frac{1}{z} H_1(z,\theta)$$

$$+ \operatorname{Res}_{z=p_1^*} \frac{1}{(z-p_1)\left(\dfrac{1}{z}-p_1\right)(z-p_1^*)\left(\dfrac{1}{z}-p_1^*\right)} \frac{1}{z} H_1(z,\theta). \tag{4.181}$$

Since by construction there is no sub-sequence of zeros such that equation (4.177) is true, we have in both cases

$$H_1(p_1,\theta) = H_1(P_1,\theta) + O\left(\frac{1}{\tau}\right) \tag{4.182}$$

and $H_1(P_1) \neq 0$.

Elaborating (4.180) gives

$$= \frac{-1}{p_1 - \dfrac{1}{p_1}} \frac{1}{p_1} H_1(p_1,\theta). \tag{4.183}$$

Using equation (4.176) it can be seen that

$$\frac{1}{\dfrac{1}{p_1} - p_1} = \frac{1}{\dfrac{1}{P_1\left(1-\frac{1}{\tau}\right)} - P_1\left(1-\dfrac{1}{\tau}\right)}. \tag{4.184}$$

Using the fact that $P_1 = \frac{1}{P_1}$ for real poles and

$$\frac{1}{1-\dfrac{1}{\tau}} - \left(1-\frac{1}{\tau}\right) = 1 + \frac{1}{\tau}\frac{1}{1-\dfrac{1}{\tau}} - 1 + \frac{1}{\tau}$$

$$= \frac{1}{\tau}\left(1 + \frac{1}{1 - \frac{1}{\tau}}\right) = \frac{1}{\tau}\left(2 + O\left(\frac{1}{\tau}\right)\right) \tag{4.185}$$

the first factor of (4.183) is

$$\frac{-1}{p_1 - \frac{1}{p_1}} = \frac{P_1}{2}\tau + O(1). \tag{4.186}$$

Since $p_1 \to \mathbb{T}$ as $\tau \to \infty$ and $H_1(z)$ has no pole in a neighborhood of P_1

$$\frac{1}{p_1}H_1(p_1, \theta) \tag{4.187}$$

is bounded. Therefore, collecting the results leads to

$$\text{Res}_{z=p_1} H(z^{-1}, \theta)z^{-1}H(z, \theta) = \tau\left(\frac{1}{2} + O\left(\frac{1}{\tau}\right)\right)H_1(p_1, \theta). \tag{4.188}$$

Applying the notation $K_{H_1} = \lim_{\tau \to \infty} H_1(p_1, \theta)$ (4.69) can be written as

$$\overline{C}(\theta, \tau) = \frac{\left(\frac{1}{\tau^\sigma}K_{T_2} + o\left(\frac{1}{\tau^\sigma}\right)\right)^2}{\tau\left(\frac{1}{2} + O\left(\frac{1}{\tau}\right)\right)\left(K_{H_1} + o(1)\right)} = \frac{1}{\tau}\frac{2\frac{1}{\tau^{2\sigma}}K_{T_{-1}}^2}{K_{H_1}} + o\left(\frac{1}{\tau}\right). \tag{4.189}$$

We can conclude that the convergence rate of the corresponding $\overline{C}(\theta, \tau)$ is $\frac{1}{\tau^{2\sigma+1}}$.

In the case of conjugate pole pairs p_1 can be written as

$$p_1 = r_{p_1}e^{j\theta_1} \tag{4.190}$$

where $r_{p_1}, \theta_1 \in \mathbb{R}$. The assumption states that $r_{p_1} = 1 - \frac{1}{\tau}$. Elaborating (4.181) leads to

$$= \left(\frac{H_1(p_1, \theta)}{\left(\frac{1}{p_1} - p_1\right)(p_1 - p_1^*)\left(\frac{1}{p_1} - p_1^*\right)}\frac{1}{p_1}\right.$$

$$+ \left.\frac{H_1(p_1^*, \theta)}{(p_1^* - p_1)\left(\frac{1}{p_1^*} - p_1\right)\left(\frac{1}{p_1^*} - p_1^*\right)}\frac{1}{p_1^*}\right)$$

$$= \frac{1}{p_1 - p_1^*}\frac{1}{1 - |p_1|^2}\left(\frac{H_1(p_1, \theta)}{\frac{1}{p_1} - p_1} - \frac{H_1(p_1^*, \theta)}{\frac{1}{p_1^*} - p_1^*}\right)$$

$$= \frac{1}{\mathrm{Im}\{p_1\}} \frac{1}{1 - |p_1|^2} \mathrm{Im}\left\{\frac{H_1(p_1)}{\frac{1}{p_1} - p_1}\right\}. \tag{4.191}$$

Let us discuss the factors of (4.191) separately. The first factor is

$$\frac{1}{\mathrm{Im}\{p_1\}} = \frac{1}{r_p \sin\theta_1} = \frac{1}{\left(1 - \frac{1}{\tau}\right)\sin\theta_1} = \frac{1}{\sin\theta_1}\left(1 + O\left(\frac{1}{\tau}\right)\right). \tag{4.192}$$

The second factor of (4.191) can be written as

$$\frac{1}{1 - |p_1|^2} = \frac{1}{1 - \left(1 - \frac{1}{\tau}\right)^2} = \frac{1}{\frac{2}{\tau} - \frac{1}{\tau^2}} = \frac{\tau}{2}\left(1 + O\left(\frac{1}{\tau}\right)\right). \tag{4.193}$$

The denominator of the third factor of (4.191) is

$$\frac{1}{p_1} - p_1 = \frac{1}{\left(1 - \frac{1}{\tau}\right)}e^{-j\theta_1} - \left(1 - \frac{1}{\tau}\right)e^{j\theta_1}$$

$$= \left(1 + \frac{1}{\tau} + O\left(\frac{1}{\tau^2}\right)\right)e^{-j\theta_1} - e^{j\theta_1} + \frac{1}{\tau}e^{j\theta_1}$$

$$= e^{-j\theta_1} - e^{j\theta_1} + \frac{1}{\tau}\left(e^{-j\theta_1} + e^{j\theta_1}\right) + O\left(\frac{1}{\tau^2}\right) = 2j\sin(-\theta_1) + O\left(\frac{1}{\tau}\right). \tag{4.194}$$

Therefore

$$\mathrm{Im}\left\{\frac{H_1(p_1,\theta)}{\frac{1}{p_1} - p_1}\right\} = \mathrm{Im}\left\{\frac{H_1(e^{j\theta_1},\theta)}{2j\sin(-\theta_1) + O\left(\frac{1}{\tau}\right)}\right\}$$

$$= -\frac{\mathrm{Re}\{H_1(e^{j\theta_1},\theta)\}}{2\sin\theta_1} + O\left(\frac{1}{\tau}\right) = -\frac{\mathrm{Re}\{H_1(e^{j\theta_1},\theta)\}}{2\sin\theta_1} + o(1). \tag{4.195}$$

Collecting the results and using the same arguments like in the single pole case we can conclude that the $\overline{C}(\theta,\tau)$ converges to zero at the rate $1/\tau$.

Here we used the fact that in the case of a complex conjugate pole pair there always exists a point on $\mathbb{T}$ such that no zero converges to it. Using the fact that the larger $|T_2(\theta,\tau)|$ the smaller is the cost function, and the first note 4.2.5 gives that there exists a τ where the set of poles (4.176) gives a smaller cost function than the set of the poles (4.173). This is a contradiction because we have assumed that the former series of poles gives for any τ

the smallest cost function.

4.4.6   Calculating Integrals via the Residue Theorem

Up to this point the assumption $n_\beta + n_\alpha \leq 2$ was not used. In this and the forthcoming sub-sections this condition is covered. This assumption is applied when the terms $T_2(\theta, \tau)$ and $T_3(\theta)$ have to be calculated. (4.51) and (4.52) show that the integrals can be determined by using the residue theorem.

4.4.6.1   Investigation of the Contribution of the Term $T_2(\theta, \tau)$

We know that

$$\left(\int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega \right)^2$$

$$= \left(2\pi \sum_k \mathrm{Res}_{z=p_k} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) + 2\pi \sum_k \mathrm{Res}_{z=p_{T,k}} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) \right)^2 \quad (4.196)$$

where for $\tau > 1$ the terms in the sum correspond to the poles $p_{T,k}$ and p_k of the target and the approximator function, respectively.

Let us introduce the constant K_T as

$$K_T = \max \{ |q_{T,1}|, \ldots, |q_{T,n_{\alpha T}}| \} \quad (4.197)$$

where

$$q_{T,m} = \begin{cases} p_{T,m}, & \text{if } p_{T,m} \in \mathbb{D}, \\ \dfrac{1}{p_{T,m}} & \text{otherwise.} \end{cases} \quad (4.198)$$

By definition $K_T < 1$. Using the definition from the previous subsection it follows that

$$= \left(\mathrm{Res}_{z=p_1} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) + O\left(K_T^\tau \tau^{n_{\alpha T}-1} \right) \right)^2 \quad (4.199)$$

when $n_\alpha = 1$ or $n_\alpha = 2$, but only one of the poles is inside the unit circle; and

$$= \left(\mathrm{Res}_{z=p_1} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) + \mathrm{Res}_{z=p_2} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) + O\left(K_T^\tau \tau^{n_{\alpha T}-1} \right) \right)^2 \quad (4.200)$$

when $n_\alpha = 2$ and both poles are inside the unit circle. From the previous section we know that the case where both p_1 and p_2 are outside the unit circle does not occur at the minimizer θ_*. The presence of the factor $\tau^{n_{\alpha T}-1}$ in (4.199) and (4.200) is motivated because of the multiple poles of the target function $T(z)$.

According to the note 4.2.2 (4.199) can be continued as

$$= \left(\mathrm{Res}_{z=p_1} T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right)^2 + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \tag{4.201}$$

and in the case of (4.200)

$$= \left(\mathrm{Res}_{z=p_1} T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta) + \mathrm{Res}_{z=p_2} T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right)^2 + O\left(K_T^\tau \tau^{n_\alpha T}\right) \tag{4.202}$$

It is worth noting that in the last case $\tau^{n_\alpha T - 1}$ was replaced by $\tau^{n_\alpha T}$. The reason for this is that $\widehat{H}(z,\theta)$ may have a double pole whose contribution can be upper bounded by the factor τ.

4.4.6.2 Investigation of the Contribution of the Term $T_3(\theta)$

This part of the chapter covers the analytic calculation of the integral $\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega$ with $n_\beta + n_\alpha \leq 2$. The results will be used when the convergence rate of the fastest poles is estimated for the minimizer θ_* and when all the poles are investigated. In the sequel we distinguish the three cases that satisfy $n_\beta + n_\alpha \leq 2$: $n_\alpha = 1$ and $n_\beta = 0$; $n_\alpha = n_\beta = 1$; and $n_\alpha = 2$ and $n_\beta = 0$.

First, we assume that $n_\alpha = 1$ and $n_\beta = 0$.

$$\frac{1}{2\pi}\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta) = \frac{1}{(z-p_1)z\left(\frac{1}{z} - p_1\right)} = \frac{-1}{(z-p_1)p_1\left(z - \frac{1}{p_1}\right)} \tag{4.203}$$

In this case $\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta)$ has only one pole inside the unit circle. Therefore

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \mathrm{Res}_{z=p_1}\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta) = \frac{-1}{p_1\left(p_1 - \frac{1}{p_1}\right)}. \tag{4.204}$$

Next, if $n_\alpha = n_\beta = 1$ then $\widehat{H}(z,\theta)$ has also a zero z_1. Evaluation of the expression $\widehat{H}(z^{-1})z^{-1}\widehat{H}(z,\theta)$ leads to

$$\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta) = \frac{(z-z_1)\left(\frac{1}{z} - z_1\right)}{(z-p_1)z\left(\frac{1}{z} - p_1\right)} = \frac{(z-z_1)z_1\left(z - \frac{1}{z_1}\right)}{(z-p_1)zp_1\left(z - \frac{1}{p_1}\right)} \tag{4.205}$$

This meromorphic function has two poles: $z = 0$ and $z = p_1$. Hence

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \mathrm{Res}_{z=0}\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta) + \mathrm{Res}_{z=p_1}\widehat{H}(z^{-1},\theta)z^{-1}\widehat{H}(z,\theta)$$

$$= \frac{z_1}{p_1} + \frac{(p_1 - z_1)z_1\left(p_1 - \frac{1}{z_1}\right)}{p_1^2\left(p_1 - \frac{1}{p_1}\right)}. \tag{4.206}$$

or

$$= \frac{z_1}{p_1} + \frac{(p_1 - z_1)\left(\frac{1}{p_1} - z_1\right)}{p_1\left(\frac{1}{p_1} - p_1\right)}. \tag{4.207}$$

Finally, the last case of this subsection is when we assume that $n_\alpha = 2$ and $n_\beta = 0$. According to the previous section, the case where $\widehat{H}(z, \theta)$ has at least one stable pole has to be covered. p_1 will be assumed to be always stable. Depending on the situations if the pole p_2 is inside or outside the unit circle different formulations of $\widehat{H}(z^{-1}, \theta)z^{-1}\widehat{H}(z, \theta)$ may be possible.

Calculating $\widehat{H}(z^{-1}, \theta)z^{-1}\widehat{H}(z, \theta)$ is divided into two cases. The first case if $p_1 \neq \frac{1}{p_2}$ or $p_1 \neq p_2$. In the second case we assume that $p_1 = p_2$ or $p_1 = \frac{1}{p_2}$.

Considering the first case:

$$\widehat{H}(z^{-1}, \theta)z^{-1}\widehat{H}(z, \theta) = \frac{1}{(z - p_1)(z - p_2)z\left(\frac{1}{z} - p_1\right)\left(\frac{1}{z} - p_2\right)}$$

$$= \frac{z}{(z - p_1)(z - p_2)p_1 p_2\left(z - \frac{1}{p_1}\right)\left(z - \frac{1}{p_2}\right)} \tag{4.208}$$

If p_2 is unstable then the analytic function (4.208) has two poles (p_1 and $\frac{1}{p_2}$) inside the unit circle. Therefore

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \text{Res}_{z=p_1}\widehat{H}(z^{-1}, \theta)z^{-1}\widehat{H}(z, \theta) + \text{Res}_{z=\frac{1}{p_2}}\widehat{H}(z^{-1}, \theta)z^{-1}\widehat{H}(z, \theta)$$

$$= \frac{1}{(p_1 - p_2)\left(p_1 - \frac{1}{p_1}\right)(p_1 p_2 - 1)} + \frac{1}{(1 - p_1 p_2)\left(\frac{1}{p_2} - p_2\right)(p_1 - p_2)}$$

$$= \frac{1}{(p_1 - p_2)(p_1 p_2 - 1)}\left(\frac{1}{p_1 - \frac{1}{p_1}} + \frac{1}{p_2 - \frac{1}{p_2}}\right)$$

$$= \frac{1}{(p_1 - p_2)(p_1 p_2 - 1)}\left(\frac{p_1}{p_1^2 - 1} + \frac{p_2}{p_2^2 - 1}\right)$$

$$= \frac{p_1(p_2^2 - 1) + p_2(p_1^2 - 1)}{(p_1 - p_2)(p_1 p_2 - 1)(p_1^2 - 1)(p_2^2 - 1)} = \frac{p_1 p_2(p_1 + p_2) - (p_1 + p_2)}{(p_1 - p_2)(p_1 p_2 - 1)(p_1^2 - 1)(p_2^2 - 1)}$$

$$= \frac{(p_1 + p_2)}{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)} \tag{4.209}$$

If p_2 is stable and $p_1 \neq p_2$ then the analytic function (4.208) has two poles (p_1 and p_2) inside the unit circle. Therefore

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \mathrm{Res}_{z=p_1} \widehat{H}(z^{-1}, \theta) z^{-1} \widehat{H}(z, \theta) + \mathrm{Res}_{z=p_2} \widehat{H}(z^{-1}, \theta) z^{-1} \widehat{H}(z, \theta)$$

$$= \frac{1}{(p_1 - p_2)\left(p_1 - \dfrac{1}{p_1}\right)(p_1 p_2 - 1)} + \frac{1}{(p_2 - p_1)\left(p_2 - \dfrac{1}{p_2}\right)(p_1 p_2 - 1)}$$

$$= \frac{1}{(p_1 - p_2)(p_1 p_2 - 1)}\left(\frac{1}{p_1 - \dfrac{1}{p_1}} - \frac{1}{p_2 - \dfrac{1}{p_2}}\right)$$

$$= \frac{1}{(p_1 - p_2)(p_1 p_2 - 1)}\left(\frac{p_1}{p_1^2 - 1} - \frac{p_2}{p_2^2 - 1}\right)$$

$$= \frac{p_1(p_2^2 - 1) - p_2(p_1^2 - 1)}{(p_1 - p_2)(p_1 p_2 - 1)(p_1^2 - 1)(p_2^2 - 1)} = \frac{-p_1 p_2(p_1 - p_2) - (p_1 - p_2)}{(p_1 - p_2)(p_1 p_2 - 1)(p_1^2 - 1)(p_2^2 - 1)}$$

$$= \frac{(1 + p_1 p_2)}{(1 - p_1 p_2)(p_1^2 - 1)(p_2^2 - 1)} \tag{4.210}$$

Consider now the second case where $p_1 = \frac{1}{p_2}$. The analytic function (4.208) simplifies to

$$= \frac{z}{(z - p_1)^2 \left(z - \dfrac{1}{p_1}\right)^2} \tag{4.211}$$

This meromorphic function has a double pole at p_1 inside the unit circle. The residue is determined after differentiation:

$$\left[\frac{z}{\left(z - \dfrac{1}{p_1}\right)^2}\right]' = \frac{\left(z - \dfrac{1}{p_1}\right)^2 - 2z\left(z - \dfrac{1}{p_1}\right)}{\left(z - \dfrac{1}{p_1}\right)^4} =$$

$$
= \frac{z - \dfrac{1}{p_1} - 2z}{\left(z - \dfrac{1}{p_1}\right)^3} = -\frac{z + \dfrac{1}{p_1}}{\left(z - \dfrac{1}{p_1}\right)^3}
\tag{4.212}
$$

Hence the contribution of the term $T_3(\theta)$ is

$$
\frac{1}{2\pi} \int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \mathrm{Res}_{z=p_1} \widehat{H}(z^{-1},\theta) z^{-1} \widehat{H}(z,\theta) = -\frac{p_1 + \dfrac{1}{p_1}}{\left(p_1 - \dfrac{1}{p_1}\right)^3}
\tag{4.213}
$$

Finally, the second case where $p_1 = p_2$ reduces equation (4.208) to

$$
\frac{1}{(z - p_1)^2 \left(\dfrac{1}{z} - p_1\right)^2 z} = \frac{z}{(z - p_1)^2 p_1^2 \left(z - \dfrac{1}{p_1}\right)^2}.
\tag{4.214}
$$

The analytic function $\widehat{H}(z^{-1},\theta) z^{-1} \widehat{H}(z,\theta)$ has a double pole at p_1 inside the unit circle. The residue at the double pole is calculated by differentiation in the neighborhood of p_1:

$$
\left[\frac{1}{\left(\dfrac{1}{z} - p_1\right)^2 z}\right]' = -\frac{-2 \left(\dfrac{1}{z} - p_1\right) \dfrac{1}{z^2} z + \left(\dfrac{1}{z} - p_1\right)^2}{\left(\dfrac{1}{z} - p_1\right)^4 z^2}
$$

$$
= -\frac{-2\dfrac{1}{z} + \dfrac{1}{z} - p_1}{\left(\dfrac{1}{z} - p_1\right)^3 z^2} = \frac{\dfrac{1}{z} + p_1}{\left(\dfrac{1}{z} - p_1\right)^3 z^2}.
\tag{4.215}
$$

Substituting $z = p_1$ it leads to

$$
\frac{\dfrac{1}{p_1} + p_1}{\left(\dfrac{1}{p_1} - p_1\right)^3 p_1^2} = \frac{1 + p_1^2}{(1 - p_1^2)^3}
\tag{4.216}
$$

The contribution of the term $T_3(\theta)$ is

$$
\frac{1}{2\pi} \int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \mathrm{Res}_{z=p_1} \widehat{H}(z^{-1}) z^{-1} \widehat{H}(z,\theta) = \frac{1 + p_1^2}{(1 - p_1^2)^3}.
\tag{4.217}
$$

4.4.7   Convergence rate

Until now it has been shown that at least one pole converges to the unit circle. Now the convergence rate of this pole or of these poles are studied. The upper and the lower

bounds of the optimal convergence rate is determined. Since these bounds are equal to the optimal converge rate in the sense that $\overline{C}(\theta_*, \tau)$ is asymptotically equivalent with $\frac{c_1}{\tau^{\sigma+1}}$ if n_α is odd and $\frac{c_1}{\tau}$ if n_α is even, where c_1 is an appropriate, non-zero constant. The optimal convergence rate is not only an interesting statement but it will be strongly used in §4.4.8. In order to simplify the treatment, the notation σ is overloaded. In the case of real poles of $H(\Omega, \theta)$ the original definition is applied. If $n_\alpha = 2$ then $n_\beta + n_\alpha \leq 2$ leads to $\sigma = 0$. Therefore, in the following sections if not stated otherwise, $\sigma = 0$ is assumed.

In this part of the chapter the notation for the members of the set which contains all the poles converging to the unit circle is

$$p_k(\tau) = P_k(1 \pm a_{k,\tau}) \tag{4.218}$$

where $P_k = \pm 1$, $a_{k,\tau} > 0$, $a_{k,\tau} \to 0$. According to note 4.2.3 and using the fact that $P_k = \frac{1}{P_k}$

$$\frac{1}{p_k(\tau)} = P_k(1 + a_{k,\tau} + O(a_{k,\tau}^2)). \tag{4.219}$$

Similarly, the zeros are denoted as $z_k = Z_k(1 - b_{k,\tau})$, where $b_{k,\tau} \to 0$. Defining $C_{a,k}$ as

$$C_{a,k} = (1 - a_{k,\tau})^\tau, \tag{4.220}$$

we will conclude after of the proof of the upper bound that the overall $\overline{C}(\theta, \tau)$ can be written as

$$\overline{C}(\theta, \tau) = O(a_{k,\tau}^{\sigma+1})C_{a,k}^2 + O(K_T^\tau \tau^{n_\alpha T - 1}). \tag{4.221}$$

To prove this upper bound we suppose that $a_{k,\tau} = O\left(\frac{1}{\tau}\right)$. If the optimal θ_* contained a pole which converges to $\mathbb{T}$ faster than $\frac{1}{\tau}$ then $C_{a,k}$ would be bounded and the overall convergence rate would be determined by the factor $O(a_{k,\tau}^{\sigma+1} + K_T^\tau \tau^{n_\alpha T - 1})$. Therefore $\overline{C}(\theta, \tau)$ would decrease faster than $\frac{1}{\tau^{\sigma+1}}$. Thus it is a contradiction showing that the optimal convergence rate cannot be faster than $\frac{1}{\tau}$.

In the cases when the lower bound is determined we suppose that $a_{k,\tau} = \frac{d_{k,\tau}}{\tau}$ where $d_{k,\tau} \leq \tau$ and $d_{k,\tau} \to \infty$. $\overline{C}(\theta, \tau)$ can be formulated as

$$\overline{C}(\theta, \tau) = O\left(\left(\frac{d_{k,\tau}}{\tau}\right)^{\sigma+1}\right) O\left(K^{d_{k,\tau}}\right) + O(K_T^\tau \tau^{n_\alpha T - 1}) \tag{4.222}$$

where $|K| < 1$ is non-zero constant. According to note 4.2.9 it follows that $\overline{C}(\theta, \tau) = o\left(\frac{1}{\tau^{\sigma+1}}\right)$. Since in the last section construction of θ with convergence rate $\frac{1}{\tau^{\sigma+1}}$ was shown, this case does not belong to the set of possible optimal parameter vectors.

Like in the previous subsection some sub-cases have to be taken into account. Depending on the orders of the polynomials and the position of the second pole, if it exists the following configurations will be investigated:

- $n_\beta = 0$, $n_\alpha = 1$.

- $n_\beta = 1$, $n_\alpha = 1$, $Z_1 \neq P_1$.

- $n_\beta = 1$, $n_\alpha = 1$, $Z_1 = P_1$.

- $n_\beta = 0$, $n_\alpha = 2$, p_1 stable, p_2 unstable and $p_1 \neq \frac{1}{p_2}$.

- $n_\beta = 0$, $n_\alpha = 2$, p_1 stable, p_2 unstable and $p_1 = \frac{1}{p_2}$.

- $n_\beta = 0$, $n_\alpha = 2$, p_1 stable, p_2 stable and $p_1 \neq p_2$.

- $n_\beta = 0$, $n_\alpha = 2$, p_1 stable, p_2 stable and $p_1 = p_2$.

Assuming that $T\left(\frac{1}{z}\right)$ has zero at $z = P$ with multiplicity m. $\tilde{T}_P(z)$ is defined as

$$\tilde{T}_P(z) = \frac{T\left(\frac{1}{z}\right)}{(z - P)^m}. \tag{4.223}$$

4.4.7.1   Case $n_\beta = 0$ and $n_\alpha = 1$

The content of the parentheses in (4.199) can be written as

$$\text{Res}_{z=p_1} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) = (\tilde{T}_{P_1}(P_1^{-1}) a_{1,\tau}^\sigma + O(a_{1,\tau}^{\sigma+1})) C_{a,1}$$

$$= \tilde{T}_{P_1}(P_1^{-1}) C_{a,1} a_{1,\tau}^\sigma + C_{a,1} O(a_{1,\tau}^{\sigma+1}) \tag{4.224}$$

Using (4.204) and (4.218) the denominator in (4.69) can be expressed as

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1}{P_1(1 - a_{1,\tau}) \left(\frac{P_1}{1 - a_{1,\tau}} - P_1 + P_1 a_{1,\tau} \right)} \tag{4.225}$$

Since

$$\left(\frac{P_1}{1 - a_{1,\tau}} - P_1 + P_1 a_{1,\tau} \right) = P_1(2a_{1,\tau} + O(a_{1,\tau}^2)) = 2P_1 a_{1,\tau}(1 + O(a_{1,\tau})) \tag{4.226}$$

we have

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1}{(1 - a_{1,\tau}) 2a_{1,\tau}(1 + O(a_{1,\tau}))}$$

$$= \frac{1}{2a_{1,\tau}} (1 + O(a_{1,\tau})) = \frac{1}{2a_{1,\tau}} + O(1). \tag{4.227}$$

Combining the results gives

$$\overline{C}(\theta, \tau) = \frac{\left(\tilde{T}_{P_1}(P_1^{-1}) C_{a,1} a_{1,\tau}^\sigma + C_{a,1} O(a_{1,\tau}^{\sigma+1}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \right)^2}{\frac{1}{2a_{1,\tau}} + O(1)}$$

$$= C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1} + a_{1,\tau}^{\sigma+1} K_T^\tau \tau^{n_\alpha T - 1}). \tag{4.228}$$

4.4.7.2 Case $n_\beta = 1$, $n_\alpha = 1$ and $Z_1 \neq P_1$

Each zero can be represented like poles in the equation (4.218):

$$z_k(\tau) = Z_k(1 + b_{k,\tau}) \tag{4.229}$$

where $|Z_k| = 1$ and $b_{k,\tau} \to 0$ as $\tau \to \infty$.

In the first case we assume that the zero does not converge to P_1, $z_1(\tau) \to Z_1 \neq P_1(\tau \to \infty)$. The content of the parentheses in (4.199) can be written as

$$\mathrm{Res}_{z=p_1} T(z^{-1}) z^{\tau-1} H_1(z, \theta)$$

$$= (\tilde{T}_{P_1}(P_1^{-1}) a_{1,\tau}^\sigma + O(a_{1,\tau}^{\sigma+1})) C_{a,1}((P_1 - Z_1) + O(a_{1,\tau} + b_{1,\tau}))$$

$$= \tilde{T}_{P_1}(P_1^{-1}) C_{a,1}(P_1 - Z_1) a_{1,\tau}^\sigma + C_{a,1} O(a_{1,\tau}^\sigma (a_{1,\tau} + |b_{1,\tau}|)) \tag{4.230}$$

Using (4.207), (4.218) and (4.229) the denominator in (4.69) can be written as

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{Z_1(1 - b_{1,\tau})}{P_1(1 - a_{1,\tau})} +$$

$$\frac{(P_1 - Z_1 + Z_1 b_{1,\tau} - P_1 a_{1,\tau}) \left(P_1 \dfrac{1}{1 - a_{1,\tau}} - Z_1 + Z_1 b_{1,\tau} \right)}{P_1(1 - a_{1,\tau}) \left(\dfrac{P_1}{1 - a_{1,\tau}} - P_1 + P_1 a_{1,\tau} \right)}$$

Since

$$\left(\frac{P_1}{1 - a_{1,\tau}} - P_1 + P_1 a_{1,\tau} \right) = P_1(2a_{1,\tau} + O(a_{1,\tau}^2)) = 2P_1 a_{1,\tau}(1 + O(a_{1,\tau})) \tag{4.231}$$

and

$$(P_1 - Z_1 + Z_1 b_{1,\tau} - P_1 a_{1,\tau}) \left(P_1 \frac{1}{1 - a_{1,\tau}} - Z_1 + Z_1 b_{1,\tau} \right)$$

$$= (P_1 - Z_1)^2 (1 + O(a_{1,\tau} + b_{1,\tau})) \tag{4.232}$$

we have

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{Z_1}{P_1} + \frac{(P_1 - Z_1)^2}{2a_{1,\tau}} (1 + O(a_{1,\tau} + |b_{1,\tau}|))$$

$$= \frac{1}{2a_{1,\tau}}(P_1 - Z_1)^2 + O(1). \tag{4.233}$$

Combining the results gives

$$\overline{C}(\theta, \tau) = \frac{\left(\tilde{T}_{P_1}(P_1^{-1})C_{a,1}(P_1 - Z_1)a_{1,\tau}^{\sigma} + C_{a,1}O(a_{1,\tau}^{\sigma}(a_{1,\tau} + |b_{1,\tau}|)) + O(K_T^{\tau}\tau^{n_{\alpha}T-1}) \right)^2}{\frac{1}{2a_{1,\tau}}(P_1 - Z_1)^2 + O(1)}$$

$$= C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1} + a_{1,\tau}^{\sigma+1}K_T^{\tau}\tau^{n_{\alpha}T-1}) \tag{4.234}$$

4.4.7.3   Case $n_{\beta} = 1$, $n_{\alpha} = 1$, and $Z_1 = P_1$

In this case the assumption $Z_1 = P_1$ means that the pole and the zero of $H_1(z, \theta)$ converge to the same point on the unit circle. According to (4.199) the following component has to be studied:

$$\text{Res}_{z=p_1} T(z^{-1})z^{\tau-1}\widehat{H}(z, \theta) = (\tilde{T}_{P_1}(P_1^{-1})a_{1,\tau}^{\sigma} + O(a_{1,\tau}^{\sigma+1}))C_{a,1}P_1(-a_{1,\tau} + b_{1,\tau}) \tag{4.235}$$

It is worth noting that it is not assumed that $b_{1,\tau} > 0$. Because of the irreducibility of the θ it is true that $\forall \tau : a_{1,\tau} \neq b_{1,\tau}$.

- First we assume that $\frac{b_{1,\tau}}{a_{1,\tau}} = O(1)$. It means that $b_{1,\tau}$ converges to 0 at least as fast as $a_{1,\tau}$. Because of the assumptions $\left(\frac{b_{1,\tau}}{a_{1,\tau}} - 1\right)$ is bounded and converges to a non-zero constant.

$$\lim_{\tau \to \infty} \left(\frac{b_{1,\tau}}{a_{1,\tau}} - 1 \right) = C_b \tag{4.236}$$

Using (4.236), (4.235) becomes

$$= a_{1,\tau}^{\sigma+1}C_b\tilde{T}_{P_1}(P_1^{-1})P_1 C_{a,1} + C_{a,1}C_b O(a_{1,\tau}^{\sigma+2}). \tag{4.237}$$

Substituting $p_1 = P_1(1 - a_{1,\tau})$ and $z_1 = P_1(1 - b_{1,\tau})$ into (4.207) gives

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1 - b_{1,\tau}}{1 - a_{1,\tau}} + \frac{P_1^2(-a_{1,\tau} + b_{1,\tau})(a_{1,\tau} + b_{1,\tau} + O(a_{1,\tau}^2))}{P_1^2(1 - a_{1,\tau})(2a_{1,\tau} + O(a_{1,\tau}^2))}. \tag{4.238}$$

Since the first term converges to 1 and

$$\frac{P_1^2(-a_{1,\tau} + b_{1,\tau})(a_{1,\tau} + b_{1,\tau} + O(a_{1,\tau}^2))}{P_1^2(1 - a_{1,\tau})(2a_{1,\tau} + O(a_{1,\tau}^2))} =$$

$$\frac{a_{1,\tau}^2 \left(-1 + \frac{b_{1,\tau}}{a_{1,\tau}}\right)\left(1 + \frac{b_{1,\tau}}{a_{1,\tau}} + O(a_{1,\tau}^2)\right)}{2a_{1,\tau}(1 - a_{1,\tau})(1 + O(a_{1,\tau}))} = O(a_{1,\tau}) \tag{4.239}$$

we conclude that $\frac{1}{2\pi}\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = 1 + O(a_{1,\tau})$. (4.238) converges to 1 as $\tau \to \infty$. Together with (4.237) it leads to

$$\overline{C}(\theta,\tau) = \frac{\left(a_{1,\tau}^{\sigma+1} C_b \tilde{T}_{P_1}(P_1^{-1}) P_1 C_{a,1} + C_{a,1} C_b O(a_{1,\tau}^{\sigma+2}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)\right)^2}{1 + O(a_{1,\tau})}$$

$$= a_{1,\tau}^{2\sigma+2} C_{a,1}^2 C_b^2 (\tilde{T}_{P_1}(P_1^{-1}))^2 + C_{a,1}^2 O(a_{1,\tau}^{2\sigma+3}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \qquad (4.240)$$

- Second $\left|\frac{b_{1,\tau}}{a_{1,\tau}}\right| \to \infty$, i.e. $\frac{a_{1,\tau}}{b_{1,\tau}} = o(1)$ is supposed. It means that a_τ converges to 0 faster than b_τ. It follows that $\frac{a_{1,\tau}^2}{b_{1,\tau}} \to 0$.

Elaborating (4.235) gives a

$$= b_{1,\tau} a_{1,\tau}^\sigma \left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right) \tilde{T}_{P_1}(P_1^{-1}) P_1 C_{a,1} + C_{a,1} O(a_{1,\tau}^{\sigma+1} b_{1,\tau}). \qquad (4.241)$$

Expression (4.238) yields

$$= \frac{1 - b_{1,\tau}}{1 - a_{1,\tau}} + \frac{b_{1,\tau}^2 \left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right)\left(1 + \frac{a_{1,\tau}}{b_{1,\tau}} + O\left(\frac{a_{1,\tau}^2}{b_{1,\tau}}\right)\right)}{(1 - a_{1,\tau}) 2 a_{1,\tau}(1 + O(a_{1,\tau}))}. \qquad (4.242)$$

Now, two sub-cases depending on that $\frac{b_{1,\tau}^2}{a_{1,\tau}} \to 0$ or not, have to be investigated. First, we assume that $\frac{a_{1,\tau}}{b_{1,\tau}^2} = O(1)$.

$$= \frac{b_{1,\tau}^2}{a_{1,\tau}} \left(\frac{a_{1,\tau}}{b_{1,\tau}^2} \frac{1 - b_{1,\tau}}{1 - a_{1,\tau}} + \frac{\left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right)\left(1 + \frac{a_{1,\tau}}{b_{1,\tau}} + O\left(\frac{a_{1,\tau}^2}{b_{1,\tau}}\right)\right)}{(1 - a_{1,\tau}) 2(1 + O(a_{1,\tau}))} \right) \qquad (4.243)$$

Since the second term in the parenthesis converges to $\frac{1}{2}$ and the first term is positive for all τ sufficiently large, (4.243) is positive for all τ sufficiently large. It means that the expression has a lower bound and hence $\overline{C}(\theta,\tau)$ (4.69) has an upper bound.

$$\overline{C}(\theta,\tau) \leq \frac{\left(b_{1,\tau} a_{1,\tau}^\sigma \left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right) \tilde{T}_{P_1}(P_1^{-1}) P_1 C_{a,1} + C_{a,1} O(a_{1,\tau}^{\sigma+1} b_{1,\tau}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)\right)^2}{\frac{b_{1,\tau}^2}{2 a_{1,\tau}}}$$

$$= 2 a_{1,\tau}^{2\sigma+1} C_{a,1}^2 \left(\left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right) \tilde{T}_{P_1}(P_1^{-1}) P_1\right)^2$$

$$+ C_{a,1}^2 O(a_{1,\tau}^{2\sigma+2}) + \frac{a_{1,\tau}}{b_{1,\tau}^2} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) = C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \quad (4.244)$$

In the second case we suppose that $\frac{b_{1,\tau}^2}{a_{1,\tau}} \to 0$. It means that the limit of (4.242) is determined by the first term, and therefore, (4.242) converges to 1. The basic assumption $\frac{b_{1,\tau}^2}{a_{1,\tau}} \to 0$ is equivalent with $b_{1,\tau}^2 = o(a_{1,\tau})$. Collecting the numerator and the denominator of $\overline{C}(\theta, \tau)$ leads to

$$\overline{C}(\theta, \tau) = \frac{\left(b_{1,\tau} a_{1,\tau}^\sigma \left(1 - \frac{a_{1,\tau}}{b_{1,\tau}}\right) \tilde{T}_{P_1}(P_1^{-1}) P_1 C_{a,1} + C_{a,1} O(a_{1,\tau}^{\sigma+1} b_{1,\tau}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)\right)^2}{1 + o(1)}$$

$$= C_{a,1}^2 O(b_{1,\tau}^2 a_{1,\tau}^{2\sigma}) + O\left(K_T^{2\tau} \tau^{2n_\alpha T - 2}\right)$$

$$= C_{a,1}^2 o(a_{1,\tau}^{2\sigma+1}) + O\left(K_T^{2\tau} \tau^{2n_\alpha T - 2}\right). \quad (4.245)$$

4.4.7.4   Case $n_\beta = 0$, $n_\alpha = 2$, p_1 stable, p_2 unstable and $p_1 \neq \dfrac{1}{p_2}$

Substituting $\widehat{H}(z, \theta) = \frac{1}{(z - p_1)(z - p_2)}$ into (4.199) the corresponding term in the parentheses is

$$\mathrm{Res}_{z = p_1} T(z^{-1}) z^{\tau - 1} \widehat{H}(z, \theta) = T(p_1^{-1}) p_1^{\tau - 1} \frac{1}{p_1 - p_2} + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \quad (4.246)$$

Expression (4.209) shows that

$$\mathrm{Res}_{z = p_1} \widehat{H}(z^{-1}, \theta) z^{-1} \widehat{H}(z, \theta) + \mathrm{Res}_{z = \frac{1}{p_2}} \widehat{H}(z^{-1}, \theta) z^{-1} \widehat{H}(z, \theta)$$

$$= \frac{(p_1 + p_2)}{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)} \quad (4.247)$$

Combining the last two equations we get

$$\overline{C}(\theta, \tau) = \frac{\left(T(p_1^{-1}) p_1^{\tau - 1} \dfrac{1}{p_1 - p_2} + O(K_T^\tau \tau^{n_\alpha T - 1})\right)^2}{\dfrac{(p_1 + p_2)}{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)}}$$

$$= \frac{\left(T(p_1^{-1}) p_1^{\tau - 1}\right)^2 (p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)(p_1 - p_2)} + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \frac{(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} T(p_1^{-1}) p_1^{\tau - 1}$$

$$+ O\left(K_T^{2\tau} \tau^{2n_\alpha T - 2}\right) \frac{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} \quad (4.248)$$

Now, we investigate the factors in the last expression. Substituting $p_1 = P_1(1 - a_{1,\tau})$ (where $P_1 = \pm 1$) into $(1 - p_1^2)$ leads to

$$(1 - p_1^2) = (1 - P_1 + P_1 a_{1,\tau})(1 + P_1 - P_1 a_{1,\tau})$$

$$= (2 + O(a_{1,\tau}))a_{1,\tau} = 2a_{1,\tau} + O(a_{1,\tau}^2). \tag{4.249}$$

If p_2 does not converge to $\mathbb{T}$ then $P_1 \neq P_2$ and $\frac{(p_2^2 - 1)}{(p_1 + p_2)(p_1 - p_2)}$, $\frac{(p_1 - p_2)(p_2^2 - 1)}{(p_1 + p_2)}$ and $\frac{(p_2^2 - 1)}{(p_1 + p_2)}$ are bounded as $\tau \to \infty$:

$$\frac{(p_2^2 - 1)}{(p_1 + p_2)(p_1 - p_2)} = \frac{(P_2^2 - 1)}{(P_1 + P_2)(P_1 - P_2)} + O(a_{1,\tau} + a_{2,\tau}), \tag{4.250}$$

$$\frac{(p_1 - p_2)(p_2^2 - 1)}{(p_1 + p_2)} = \frac{(P_1 - P_2)(P_2^2 - 1)}{(P_1 + P_2)} + O(a_{1,\tau} + a_{2,\tau}), \tag{4.251}$$

and

$$\frac{(p_2^2 - 1)}{(p_1 + p_2)} = \frac{(P_2^2 - 1)}{(P_1 + P_2)} + O(a_{1,\tau} + a_{2,\tau}). \tag{4.252}$$

Therefore, (4.248) can be written as

$$\overline{C}(\theta, \tau) = -\frac{2a_{1,\tau}\left((T(p_1^{-1}))^2(P_2^2 - 1)\right)C_{a,1}^2}{(P_1^2 - P_2^2)}$$

$$+ C_{a,1}^2 O((a_{1,\tau} + a_{2,\tau})a_{1,\tau}^{2\sigma+1}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \overset{P_1^2 = 1}{=} 2a_{1,\tau}^{2\sigma+1}\left(\tilde{T}_{P_1}(P_1^{-1})\right)^2 C_{a,1}^2$$

$$+ C_{a,1}^2 O((a_{1,\tau} + a_{2,\tau})a_{1,\tau}^{2\sigma+1}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \tag{4.253}$$

If p_2 converges to $\mathbb{T}$ ($p_1 = P_1(1 - a_{1,\tau})$ and $p_2 = P_2(1 + a_{2,\tau})$ with $a_{1,\tau} > 0$, $a_{1,\tau} > 0$ for τ sufficiently large) then the effect of the factor $(1 - p_2^2)$ has to be studied. Depending on whether $P_1 = P_2$ or $P_1 \neq P_2$ the factor $\frac{1}{(p_1 + p_2)}$ or the factor $\frac{1}{(p_1 - p_2)}$ is bounded.

$$1 - p_2^2 = 1 - (1 + a_{2,\tau})^2 = -2a_{2,\tau} + O(a_{2,\tau}^2). \tag{4.254}$$

If $P_1 = P_2$ then using the definitions $p_1 = P_1(1 - a_{1,\tau})$ and $p_2 = P_2(1 - a_{2,\tau})$ leads to

$$p_1 + p_2 = P_1(1 - a_{1,\tau} + 1 + a_{2,\tau}) = 2P_1 + O(a_{1,\tau} + a_{2,\tau}) \tag{4.255}$$

and

$$p_1 - p_2 = P_1(-a_{1,\tau} - a_{2,\tau}) = -P_1(a_{1,\tau} + a_{2,\tau}). \tag{4.256}$$

In this case the second and third terms in (4.248) remains bounded as $\tau \to \infty$:

$$O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \frac{(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} \text{ and } O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \frac{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} \quad (4.257)$$

and the first term in (4.248) can be written as

$$\left(T(p_1^{-1})\right)^2 C_{a,1}^2 \frac{(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)(p_1 - p_2)}$$

$$= \left(T(p_1^{-1})\right)^2 C_{a,1}^2 \frac{(2a_{1,\tau} + O(a_{1,\tau}^2))(-2a_{2,\tau} + O(a_{2,\tau}^2))}{(2P_1 + O(a_{1,\tau} + a_{2,\tau}))(-P_1(a_{1,\tau} + a_{2,\tau}))}$$

$$= 2\frac{a_{1,\tau} a_{2,\tau}}{(a_{1,\tau} + a_{2,\tau})} \left(T(p_1^{-1})\right)^2 C_{a,1}^2 + O((a_{1,\tau} + a_{2,\tau})^2). \quad (4.258)$$

Assuming that, for example, $a_{1,\tau}$ converges to 0 not faster than $a_{2,\tau}$, i.e. $\frac{a_{2,\tau}}{a_{1,\tau}} = O(1)$

$$\frac{a_{1,\tau} a_{2,\tau}}{(a_{1,\tau} + a_{2,\tau})} = \frac{a_{2,\tau}}{1 + \frac{a_{2,\tau}}{a_{1,\tau}}} \sim a_{2,\tau} \quad (4.259)$$

Consequently

$$\overline{C}(\theta, \tau) = C_{a,1}^2 O(a_{2,\tau} a_{1,\tau}^{2\sigma}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \quad (4.260)$$

If $P_1 \neq P_2$ then, $P_1 = -P_2$ and

$$p_1 + p_2 = P_1(1 - a_{1,\tau} - 1 - a_{2,\tau}) = -P_1(a_{1,\tau} + a_{2,\tau}) \quad (4.261)$$

and

$$p_1 - p_2 = P_1(2 - a_{1,\tau} - a_{2,\tau}) = 2P_1 + O(a_{1,\tau} + a_{2,\tau}) \quad (4.262)$$

The first term in (4.248) is the same like in the previous case and, hence, only the second and third terms has to be investigated:

$$\frac{(p_1 - p_2)(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)$$

$$= \frac{(2P_1 + O(a_{1,\tau} + a_{2,\tau}))(2a_{1,\tau} + O(a_{1,\tau}^2))(-2a_{2,\tau} + O(a_{2,\tau}^2))}{(-P_1(a_{1,\tau} + a_{2,\tau}))} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)$$

$$= \left(8\frac{a_{1,\tau} a_{2,\tau}}{a_{1,\tau} + a_{2,\tau}} + O((a_{1,\tau} + a_{2,\tau})^2)\right) O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \quad (4.263)$$

and

$$\frac{(p_1^2 - 1)(p_2^2 - 1)}{(p_1 + p_2)} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)$$

$$= \frac{(2a_{1,\tau} + O(a_{1,\tau}^2))(-2a_{2,\tau} + O(a_{2,\tau}^2))}{(-P_1(a_{1,\tau} + a_{2,\tau}))} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right)$$

$$= \left(4P_1 \frac{a_{1,\tau} a_{2,\tau}}{a_{1,\tau} + a_{2,\tau}} + O((a_{1,\tau} + a_{2,\tau})^2)\right) O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \tag{4.264}$$

Following the same lines as in the previous case it can be calculated that this sequence is bounded. Let us assume again that $a_{1,\tau}$ converges to 0 not faster than $a_{2,\tau}$. Consequently

$$\overline{C}(\theta, \tau) = C_{a,1}^2 O(a_{2,\tau} a_{1,\tau}^{2\sigma}) + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \tag{4.265}$$

4.4.7.5 Case $n_\alpha = 2$, p_1 stable, p_2 unstable, and $p_1 = \dfrac{1}{p_2}$

This case is similar to the previous case. Here, we do not need to consider sub-cases because the assumption implicitly contains that p_2 converges to the same point of $\mathbb{T}$. Substituting $p_1 = \frac{1}{p_2}$ into the first term in the parentheses in equation (4.199)

$$\mathrm{Res}_{z=p_1} T(z^{-1}) z^{\tau - 1} \widehat{H}(z, \theta) = T(p_1^{-1}) p_1^{\tau - 1} \frac{1}{p_1 - \dfrac{1}{p_1}} + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \tag{4.266}$$

The contribution of the term $T_3(\theta)$ in $\overline{C}(\theta, \tau)$ leads to (4.213):

$$\mathrm{Res}_{z=p_1} \widehat{H}(z^{-1}, \theta) z^{-1} \widehat{H}(z, \theta) = -\frac{p_1 + \dfrac{1}{p_1}}{\left(p_1 - \dfrac{1}{p_1}\right)^3} \tag{4.267}$$

Therefore, the sequence $\overline{C}(\theta, \tau)$ (4.69) can be written as

$$\overline{C}(\theta, \tau) = -\frac{(T(p_1^{-1})^2) C_{a,1}^2 \left(p_1 - \dfrac{1}{p_1}\right)}{p_1 + \dfrac{1}{p_1}} + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) \frac{\left(p_1 - \dfrac{1}{p_1}\right)^2}{p_1 + \dfrac{1}{p_1}}$$

$$= (T(p_1^{-1})^2) C_{a,1}^2 \frac{1 - p_1^2}{1 + p_1^2} + \frac{(p_1^2 - 1)^2}{p_1(p_1^2 + 1)} O\left(K_T^\tau \tau^{n_\alpha T - 1}\right). \tag{4.268}$$

Substituting $p_1 = P_1(1 - a_{1,\tau})$ the last expression can be elaborated as

$$= (T(p_1^{-1})^2) C_{a,1}^2 \frac{2a_{1,\tau} + O(a_{1,\tau}^2)}{2 + O(a_{1,\tau})} + O\left(K_T^\tau \tau^{n_\alpha T - 1} a_{1,\tau}^2\right)$$

$$= (T(p_1^{-1})^2)C_{a,1}^2 a_{1,\tau} + (T(p_1^{-1})^2)C_{a,1}^2 O(a_{1,\tau}^2) + O\left(K_T^\tau \tau^{n_\alpha T -1} a_{1,\tau}^2\right)$$

$$= C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1}) + O(a_{1,\tau}^{2\sigma+2}) + O\left(K_T^\tau \tau^{n_\alpha T -1} a_{1,\tau}^2\right). \tag{4.269}$$

4.4.7.6 Case $n_\alpha = 2$, p_1 stable, p_2 stable and $p_1 \neq p_2$

Without loss of generality we can suppose that the pole p_1 converges to $\mathbb{T}$ at least as fast as the pole p_2, i.e. $\frac{a_{1,\tau}}{a_{2,\tau}} \to 0$ as $\tau \to \infty$. Two main sub-cases are considered: $P_1 \neq P_2$ or $P_1 = P_2$.

First, the case $P_1 \neq P_2$ is studied. The contribution of the term $T_2(\theta, \tau)$ is

$$\frac{1}{2\pi} \int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega = \frac{T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}}{p_1 - p_2} + O\left(K_T^\tau \tau^{n_\alpha T -1}\right) \tag{4.270}$$

The assumptions imply that

$$p_1 - p_2 = P_1 - P_2 + O(a_{1,\tau} + a_{2,\tau}). \tag{4.271}$$

If $|P_2| < 1$ then for all τ sufficiently large $|T(p_2^{-1})p_2^{\tau-1}| < K_2^\tau$ where $K_2 = |P_2| + \varepsilon$. If $p_2 \to \mathbb{T}$ (i.e. $p_2 \to -P_1$) then $\theta_*(\tau')$ contains only the even ($\forall x \in \tau' : \exists k \in \mathbb{Z}$ such that $x = 2k$) or the odd ($\forall x \in \tau' : \exists k \in \mathbb{Z}$ such that $x = 2k+1$) sub-sequence depending on the sign of $T(-P_1)$. Hence if θ_* represents an optimal parameter vector then

$$\left| \frac{1}{2\pi} \mathrm{Re} \left\{ \int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega \right\} \right| = \frac{|T(p_1^{-1})p_1^{\tau-1}| + |T(p_2^{-1})p_2^{\tau-1}|}{2}$$

$$+ (|T(p_1^{-1})p_1^{\tau-1}| + |T(p_2^{-1})p_2^{\tau-1}|)O(a_{1,\tau} + a_{2,\tau}) + O\left(K_T^\tau \tau^{n_\alpha T -1}\right). \tag{4.272}$$

According to (4.210) the contribution of the term $T_3(\theta)$ is

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{(1 + p_1 p_2)}{(1 - p_1 p_2)(p_1^2 - 1)(p_2^2 - 1)}. \tag{4.273}$$

This case has to be split into three sub-cases:

- $|P_2| < 1$, i.e. the stable pole p_2 does not converge to the unit circle.

- $|P_2| = 1$, but $P_2 \neq P_1$, i.e. the contribution of the residue at the point p_2 is "independent" from the contribution of the residue at the point p_1.

- $P_1 = P_2$, i.e. the contribution of the poles is similar to the contribution of a double pole.

If $|P_2| < K_2 < 1$ then only the pole p_1 converges to the unit circle. Therefore, only the convergence rate of the contribution of the sequence $p_1(\tau)$ has to be studied. The factors of (4.273) are

$$1 + p_1 p_2 = (1 + P_1 P_2) + O(a_{1,\tau} + a_{2,\tau})$$

$$1 - p_1 p_2 = (1 - P_1 P_2) + O(a_{1,\tau} + a_{2,\tau}) \tag{4.274}$$

and

$$p_2^2 - 1 = P_2^2 - 1 + O(a_{2,\tau}). \tag{4.275}$$

- The assumption $|P_2| < 1$ implies that $1 + P_1 P_2 \neq 0$, $1 - P_1 P_2 \neq 0$ and $P_2^2 - 1 \neq 0$. The contribution of the factor $p_1^2 - 1$ is

$$p_1^2 - 1 = (1 - a_{1,\tau})^2 - 1 = -2a_{1,\tau} + O(a_{1,\tau}^2). \tag{4.276}$$

Therefore,

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1}{a_{1,\tau}} \frac{(1 + P_1 P_2)}{2(1 - P_1 P_2)(1 - P_2)} + O(1). \tag{4.277}$$

It is worth noting that $1 + P_1 P_2 > 0$, $1 - P_1 P_2 > 0$ and $1 - P_2 > 0$. Hence $\overline{C}(\theta, \tau)$ (4.69) can be written as

$$\overline{C}(\theta, \tau) = \frac{(T(p_1^{-1}))^2 C_{a,1}^2}{(P_1 - P_2)^2} a_{1,\tau} \frac{2(1 - P_1 P_2)(1 - P_2)}{(1 + P_1 P_2)}$$

$$+ O(a_{1,\tau}^2 + a_{1,\tau} K_2^\tau) + |T(p_1^{-1}) p_1^{\tau-1}| O\left(a_{1,\tau} K_T^\tau \tau^{n_\alpha T - 1}\right)$$

$$= C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1}) + O(a_{1,\tau}^{2\sigma+2}). \tag{4.278}$$

- If $|P_2| = 1$, but $P_1 \neq P_2$ then the factors in (4.273) are

$$p_1^2 - 1 = (1 - a_{1,\tau})^2 - 1 = -2a_{1,\tau} + O(a_{1,\tau}^2),$$

$$p_2^2 - 1 = (1 - a_{2,\tau})^2 - 1 = -2a_{2,\tau} + O(a_{2,\tau}^2), \tag{4.279}$$

and using the fact that $P_1 P_2 = -1$ we get

$$1 + p_1 p_2 = a_{1,\tau} + a_{2,\tau} - a_{1,\tau} a_{2,\tau},$$

$$1 - p_1 p_2 = 2 + O(a_{1,\tau} + a_{2,\tau}). \tag{4.280}$$

Hence, (4.273) can be written as

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1}{8} \frac{a_{1,\tau} + a_{2,\tau} - a_{1,\tau} a_{2,\tau}}{a_{1,\tau} a_{2,\tau}} + O(a_{1,\tau} + a_{2,\tau})$$

$$= \frac{1}{8}\left(\frac{1}{a_{1,\tau}} + \frac{1}{a_{2,\tau}} - 1\right) + O(a_{1,\tau} + a_{2,\tau}). \tag{4.281}$$

Combining (4.281) and (4.272) yields

$$\overline{C}(\theta, \tau) = 2\frac{a_{1,\tau}a_{2,\tau}}{a_{1,\tau} + a_{2,\tau}}\left(|T(p_1^{-1})p_1^{\tau-1}| + |T(p_2^{-1})p_2^{\tau-1}|\right)^2$$

$$+ \left(|T(p_1^{-1})p_1^{\tau-1}| + |T(p_2^{-1})p_2^{\tau-1}|\right)^2 O((a_{1,\tau} + a_{2,\tau})a_{1,\tau}a_{2,\tau})$$

$$+ O\left(K_T^\tau\tau^{n_\alpha T - 1}\right). \tag{4.282}$$

- In the last case we assume that $P_2 = P_1$, i.e. the stable poles p_1 and p_2 converge to the same point on the unit circle. According to note 4.2.14 the term $T_2(\theta, \tau)$ can be formulated as

$$\frac{1}{2\pi}\mathrm{Re}\left\{\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}(\Omega, \theta)d\omega\right\}$$

$$= \frac{T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}}{p_1 - p_2} + O\left(K_T^\tau\tau^{n_\alpha T - 1}\right)$$

$$= (\tau - 1)T(p_*^{-1})p_*^{\tau-2} + O(1) = \tau T(p_*^{-1})p_*^{\tau-2} + O(1) \tag{4.283}$$

for some $p_* \in (p_1, p_2)$ and where in the last step we used the fact that $[T(p_*^{-1})]'p_*^{\tau-1}$ and $O\left(K_T^\tau\tau^{n_\alpha T - 1}\right)$ are bounded (see note 4.2.8). The contribution of the term $T_3(\theta)$ is (4.273). Subset of the factors were already computed in (4.279). In this case the following ones have to be calculated:

$$1 + p_1 p_2 = 2 + O(a_{1,\tau} + a_{2,\tau}),$$

$$1 - p_1 p_2 = a_{1,\tau} + a_{2,\tau} - a_{1,\tau}a_{2,\tau}. \tag{4.284}$$

Therefore,

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1}{2}\frac{1}{(a_{1,\tau} + a_{2,\tau} - a_{1,\tau}a_{2,\tau})a_{1,\tau}a_{2,\tau}} + O\left(\frac{1}{a_{1,\tau}} + \frac{1}{a_{2,\tau}}\right) \tag{4.285}$$

Collecting the results gives

$$\overline{C}(\theta, \tau) = 2(a_{1,\tau} + a_{2,\tau} - a_{1,\tau}a_{2,\tau})a_{1,\tau}a_{2,\tau}\left(\frac{T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}}{p_1 - p_2}\right)^2$$

$$+ O\left((a_{1,\tau} + a_{2,\tau} + a_{1,\tau}a_{2,\tau})\left(K_T^\tau \tau^{n_\alpha T - 1}\right)\right). \tag{4.286}$$

Without loss of generality we can suppose that $a_{1,\tau}$ converges to zero not slower than $a_{2,\tau}$, i.e. $\frac{a_{1,\tau}}{a_{2,\tau}} = O(1)$. This case has to be split into two sub-cases:

– The difference between convergence rates of $a_{1,\tau}$ and $a_{2,\tau}$ is only a constant factor, i.e. $\frac{a_{2,\tau}}{a_{1,\tau}} = O(1)$. We can conclude that $\frac{a_{1,\tau}}{a_{2,\tau}}$ converges to a nonzero constant. Since in (4.283) p_* is calculated via linear combination of p_1 and p_2, it can be written as $p_* = 1 - a_{*,\tau}$ where $a_{*,\tau} \to 0$ and its convergence rate equals to the convergence rate of $a_{1,\tau}$ or $a_{2,\tau}$.

$$\overline{C}(\theta, \tau) = 2(a_{1,\tau} + a_{2,\tau} + a_{1,\tau}a_{2,\tau})a_{1,\tau}a_{2,\tau}a_{*,\tau}^{2\sigma}\left(\tilde{T}_{P_1}(P_1^{-1})\right)^2 C_*^2 \tau^2$$

$$O\left((a_{1,\tau} + a_{2,\tau} + a_{1,\tau}a_{2,\tau})\left(K_T^\tau \tau^{n_\alpha T - 1}\right)\right). \tag{4.287}$$

– The sequence $a_{2,\tau}$ decreases slower than $a_{1,\tau}$, i.e. $\frac{a_{1,\tau}}{a_{2,\tau}} = o(1)$. From the assumptions it follows that

$$\left|p_1^{\tau-1}\right| \geq \left|p_2^{\tau-1}\right| \quad \Rightarrow \quad \left|p_1^{\tau-1} - p_2^{\tau-1}\right| \leq \left|\left|p_1^{\tau-1}\right| + \left|p_2^{\tau-1}\right|\right| \leq 2\left|p_1^{\tau-1}\right| \tag{4.288}$$

Using the last estimation the contribution of the term T_2 can be upper-bounded as

$$\left|\frac{T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}}{p_1 - p_2}\right| = \left|\frac{T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}}{-a_{1,\tau} + a_{2,\tau}}\right|$$

$$\leq \frac{2\left|p_1^{\tau-1}\right|}{a_{2,\tau}\left(1 - \frac{a_{1,\tau}}{a_{2,\tau}}\right)} \leq \frac{2\left|p_1^{\tau-1}\right|}{a_{2,\tau}} \tag{4.289}$$

Finally, the estimation w.r.t. $\overline{C}(\theta, \tau)$

$$\overline{C}(\theta, \tau) \leq 2(a_{1,\tau} + a_{2,\tau} + a_{1,\tau}a_{2,\tau})a_{1,\tau}a_{2,\tau}\left(\frac{2\left|p_1^{\tau-1}\right|}{a_{2,\tau}}\right)^2$$

$$= C_{a,1}^2 O(a_{1,\tau}^{2\sigma+1}). \tag{4.290}$$

4.4.7.7 Case $n_\alpha = 2$, p_1 stable, p_2 stable and $p_1 = p_2$

Now, we assume that $H(z, \theta)$ has a double pole at p_1. The contribution of the term $T_2(\theta, \tau)$ is the residue at the point p_1 which can be written as

$$(\tau - 1)p_1^{\tau-2}T(p_1^{-1}) + [T(p_1^{-1})]'p_1^{\tau-1} + O\left(K_T^\tau \tau^{n_\alpha T - 1}\right) = \tau p_1^{\tau-2}T(p_1^{-1}) + O(1). \tag{4.291}$$

Using the result (4.217) we get

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{1 + p_1^2}{(1 - p_1^2)^3}. \tag{4.292}$$

As we already have seen

$$1 + p_1^2 = 1 + (1 - a_{1,\tau})^2 = 2 + O(a_{1,\tau}) \tag{4.293}$$

and

$$1 - p_1^2 = 2a_{1,\tau} + O(a_{1,\tau}^2). \tag{4.294}$$

The overall contribution of the factors to $\overline{C}(\theta, \tau)$ (4.69) is

$$\overline{C}(\theta, \tau) = \frac{\left(2a_{1,\tau} + O(a_{1,\tau}^2)\right)^3 \left(\tau p_1^{\tau-2} T(p_1^{-1}) + O(1)\right)^2}{2 + O(a_{1,\tau})}$$

$$= 4a_{1,\tau}^{3+2\sigma} \tau^2 C_{a,1}^2 (\tilde{T}_{P_1}(P_1^{-1}))^2 + C_{a,1}^2 O\left(a_{1,\tau}^{3+\sigma} \tau\right) \tag{4.295}$$

q.e.d.

4.4.8   Investigation of All the Poles

In this subsection the basic assumption is that there is at least one stable pole which converges to the unit circle. In the previous subsection it is proved that the maximal convergence rate of the poles converging to $\mathbb{T}$ is $\frac{1}{\tau}$ and that θ_* contains at least one such pole. The usual notation is applied:

$$p_1(\tau) = P_1 (1 - a_{1,\tau}) \tag{4.296}$$

where $|P_1| = 1$, $a_{1,\tau} > 0$. According to the results of the previous subsection $a_{1,\tau} = \frac{c_1}{\tau}$ is supposed where $c_1 > 0$ does not depend on τ. Since the polynomials of the numerator and denominator of $H(z, \theta)$ are real, if $P_1 \in \mathbb{C}\backslash\mathbb{R}$ then there exists a $P_2 \in \mathbb{C}\backslash\mathbb{R}$, $P_1 = P_2^*$ and $p_2(\tau) = p_1^*(\tau)$. Note that the cases where complex conjugate poles occur are already covered in 4.4.5. Here the cases where p_1 is stable and p_2 is unstable are investigated. In the previous subsection the maximal convergence rate of $\overline{C}(\theta, \tau)$ is determined. Let us assume that $n_\alpha = 2$ and the poles of $\widehat{H}(z, \theta)$ are real. In this subsection it will be proven that for all τ large enough $\theta_*(\tau)$ does not contain any unstable pole.

First, the case when $\sigma \neq 0$ is covered. Assuming that $\sigma > 0$ and $\widehat{H}(z, \theta_*)$ real poles the computation of the previous subsection shows that $\overline{C}(\theta, \tau) \sim \frac{1}{\tau^{\sigma+1}}$. Since $T(z)$ has only finite number of poles, there always exists a $Q \in \mathbb{T}$ such that $T(Q) \neq 0$ and $T(Q^*) \neq 0$. The parameter vector θ which is constructed by using the following poles

$$p_1 = \left(1 - \frac{1}{\tau}\right) Q \text{ and } p_1 = \left(1 - \frac{1}{\tau}\right) Q^*, \tag{4.297}$$

results in $\overline{C}(\theta, \tau) \sim \frac{1}{\tau}$. It is a contradiction with $\theta_*(\tau)$ being global extreme value, therefore we conclude that in the case θ_* has at least two real poles, $\sigma > 0$ never occurs.

It will be proved that if $p_2(\tau)$ is an unstable pole sequence corresponds to θ, then a new θ_2 can be constructed such that $C(\theta_2, \tau) < C(\theta, \tau)$, or equivalently $\overline{C}(\theta_2, \tau) > \overline{C}(\theta, \tau)$, where θ_2 determines a stable system. It means that $p_1(\tau) = P_1 \left(1 - \frac{c_1}{\tau}\right)$ where $P_1 = \pm 1$ and that $p_2(\tau)$ cannot converge faster to $\mathbb{T}$. If the convergence rate of $p_2(\tau)$ attains its maximum then

$$p_2(\tau) = P_2 \left(1 + \frac{c_2}{\tau}\right). \tag{4.298}$$

Further in this subsection we assume that $p_1(\tau)$ and hence $p_2(\tau)$ are real poles. The investigation can be split into four different cases:

- $p_2(\tau)$ converges to $\mathbb{T}$, but $P_1 \neq P_2$.

- $p_2(\tau)$ converges to $\mathbb{T}$ and $p_1(\tau) = \frac{1}{p_2(\tau)}$.

- $p_2(\tau)$ converges to $\mathbb{T}$ and $P_1 = P_2$, but $p_1(\tau) \neq \frac{1}{p_2(\tau)}$.

- $p_2(\tau)$ does not approach to $\mathbb{T}$, i.e. $\exists \varepsilon > 0 : \forall \tau : |p_2(\tau)| \leq 1 - \varepsilon < 1$.

The technique used in this section is the following. Assuming that we have a θ that contains an unstable $p_2(\tau)$ then a θ_2 that has only stable poles is constructed such that it gives asymptotically ($\tau \to \infty$) a smaller cost function. In every case $\overline{C}(\theta, \tau)$ and $\overline{C}(\theta_2, \tau)$ are expressed as

$$\overline{C}(\theta, \tau) = \frac{1}{\tau} K_1(c_1, c_2) + o\left(\frac{1}{\tau}\right)$$

$$\overline{C}(\theta_2, \tau) = \frac{1}{\tau} K_2(c_1, c_2) + o\left(\frac{1}{\tau}\right), \tag{4.299}$$

where $K_1(c_1, c_2) > 0$ and $K_2(c_1, c_2) > 0$. The construction of θ_2 is such that

$$\max_{c_1, c_2} K_1(c_1, c_2) < \max_{c_1, c_2} K_2(c_1, c_2). \tag{4.300}$$

It proves that for every target system $T(z)$, and for every parameter vector sequence $\theta(\tau)$ representing an unstable system, a parameter vector sequence $\theta_2(\tau)$ representing a stable system can be asymptotically constructed. It means nothing else than with $n_\beta + n_\alpha \leq 2$ for all target system $T(z)$ there exists a τ such that the global minimum of the cost function (4.1) determines a stable system.

4.4.8.1 Case $|p_2(\tau)| \to 1$, with $P_1 \neq P_2$

First the denominator of (4.69) is studied. Let $a_{2,\tau}$ be a positive sequence which converges to 0.

First, we assume that $p_2(\tau) = P_2(1 - a_{2,\tau})$, i.e. p_2 is stable for τ. This is compared with the case where p_2 is unstable. According to (4.210), the contribution to $T_3(\theta)$ can be written as

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{(1 + p_1 p_2)}{(1 - p_1 p_2)(1 - p_1^2)(1 - p_2^2)} \tag{4.301}$$

Since $P_1 \neq P_2$, $|P_1| = 1$, and $|P_2| = 1$, $P_1, P_2 \in \mathbb{R}$, and $P_1 P_2 = -1$ it can be easily verified that

$$1 - p_1 p_2 = 2 - a_{1,\tau} - a_{2,\tau} + a_{1,\tau} a_{2,\tau} = 2 + O(a_{1,\tau} + a_{2,\tau}). \tag{4.302}$$

$$1 + p_1 p_2 = a_{1,\tau} + a_{2,\tau} - a_{1,\tau} a_{2,\tau} \tag{4.303}$$

Assuming that $a_{2,\tau} = \frac{c_2}{\tau}$ we get

$$1 + p_1 p_2 = \frac{c_1}{\tau} + \frac{c_2}{\tau} + O\left(\frac{1}{\tau^2}\right). \tag{4.304}$$

And if $a_{2,\tau}$ converges slower than $a_{1,\tau}$, i.e. $\frac{a_{1,\tau}}{a_{2,\tau}} = o(1)$, then

$$1 + p_1 p_2 = a_{2,\tau} + o(a_{2,\tau}). \tag{4.305}$$

$$(1 - p_1^2) = 1 - (1 - a_{1,\tau})^2 = 2a_{1,\tau} - a_{1,\tau}^2$$

$$= a_{1,\tau}(2 + O(a_{1,\tau})) = \frac{c_1}{\tau}\left(2 + O\left(\frac{1}{\tau}\right)\right) \tag{4.306}$$

Similarly

$$(1 - p_2^2) = 1 - (1 - a_{2,\tau})^2 = 2a_{2,\tau} - a_{2,\tau}^2$$

$$= a_{2,\tau}(2 + O(a_{2,\tau})) \stackrel{\text{if } a_{2,\tau} = c_2/\tau}{=} \frac{c_2}{\tau}\left(2 + O\left(\frac{1}{\tau}\right)\right). \tag{4.307}$$

Collecting the results if $a_{2,\tau} = \frac{c_2}{\tau}$ (4.301) can be written as

$$= \tau \frac{c_1 + c_2 + O\left(\dfrac{1}{\tau}\right)}{\left(2 + O\left(\dfrac{1}{\tau}\right)\right) c_1 \left(2 + O\left(\dfrac{1}{\tau}\right)\right) c_2 \left(2 + O\left(\dfrac{1}{\tau}\right)\right)} = \tau \frac{c_1 + c_2}{8 c_1 c_2} + O(1). \tag{4.308}$$

and, if $a_{2,\tau}$ is slower than $a_{1,\tau}$, then (4.301) yields

$$= \tau \frac{(a_{2,\tau} + o(a_{2,\tau}))}{(2 + O(a_{2,\tau}))c_1 \left(2 + O\left(\frac{1}{\tau}\right)\right) a_{2,\tau}(2 + O(a_{2,\tau}))} = \tau \frac{1}{8c_1} + o(\tau). \qquad (4.309)$$

Second, we assume that p_2 is unstable: $p_2(\tau) = P_2 \left(\frac{1}{1-a_{2,\tau}}\right)$ with

$$\frac{1}{1 - a_{2,\tau}} = 1 + a_{2,\tau} + O(a_{2,\tau}^2) = 1 + \frac{c_2}{\tau} + O\left(\frac{1}{\tau^2}\right). \qquad (4.310)$$

From the (4.209), the contribution of the term $T_3(\theta)$ yields

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \frac{(p_1 + p_2)}{(p_1 - p_2)(1 - p_1^2)(1 - p_2^2)} \qquad (4.311)$$

Since $P_1 \neq P_2$, $|P_1| = 1$, $|P_2| = 1$, $P_1, P_2 \in \mathbb{R}$ and $P_1 + P_2 = 0$ it follows that

$$p_1 + p_2 = P_1(1 - a_{1,\tau}) + P_2 \left(\frac{1}{1 - a_{2,\tau}}\right) =$$

$$- P_1 a_{1,\tau} + P_2 a_{2,\tau} + O(a_{2,\tau}^2) \overset{P_1 P_2 = -1}{=} -P_1(a_{1,\tau} + a_{2,\tau}) + O(a_{2,\tau}^2)$$

If $a_{2,\tau} = \frac{c_2}{\tau}$ then

$$= -\frac{P_1}{\tau}(c_1 + c_2) + O\left(\frac{1}{\tau^2}\right) \qquad (4.312)$$

and if $a_{2,\tau}$ converges slower to the unit circle than $a_{1,\tau}$ then it leads

$$= -P_1 a_{2,\tau} + O(a_{2,\tau}^2). \qquad (4.313)$$

The first factor of the denominator in (4.311) is

$$p_1 - p_2 = P_1(1 - a_{1,\tau}) + P_2 \left(\frac{1}{1 - a_{2,\tau}}\right) =$$

$$2P_1 - P_1 a_{1,\tau} + P_2 a_{2,\tau} + O(a_{2,\tau}) = P_1(2 + O(a_{1,\tau} + a_{2,\tau})) \qquad (4.314)$$

and the third factor is

$$(1 - p_2^2) = 1 - \left(1 + a_{2,\tau} + O(a_{2,\tau}^2)\right)^2$$

$$= -2a_{2,\tau} - a_{2,\tau}^2 + O(a_{2,\tau}^2) = a_{2,\tau}(-2 + O(a_{2,\tau})) \qquad (4.315)$$

which we can elaborate if $a_{2,\tau} = \frac{c_2}{\tau}$

$$1 - p_2^2 = \frac{c_2}{\tau}\left(-2 + O\left(\frac{1}{\tau}\right)\right). \tag{4.316}$$

Assuming that $a_{2,\tau} = \frac{c_2}{\tau}$ we collect the results

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \tau\frac{-P_1(c_1 + c_2) + O\left(\frac{1}{\tau}\right)}{\left(2P_1 + O\left(\frac{1}{\tau}\right)\right)c_1\left(2 + O\left(\frac{1}{\tau}\right)\right)c_2\left(-2 + O\left(\frac{1}{\tau}\right)\right)}$$

$$= \tau\frac{c_1 + c_2}{8c_1 c_2} + O(1). \tag{4.317}$$

and supposing that $a_{2,\tau}$ converges slower to the unit circle than $a_{1,\tau}$ the contribution of the term $T_3(\theta)$ can be written as

$$\frac{1}{2\pi}\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \tau\frac{(-P_1 a_{2,\tau} + O(a_{2,\tau}^2))}{P_1(2 + O(a_{2,\tau}))c_1\left(2 + O\left(\frac{1}{\tau}\right)\right)a_{2,\tau}(-2 + O(a_{2,\tau}))}$$

$$= \tau\frac{1}{8c_1} + O(\tau a_{2,\tau}). \tag{4.318}$$

Comparing (4.308) to (4.317) and (4.309) to (4.318) we see that asymptotically the denominators in (4.69) are the same. However, the numerators are different. If p_2 is unstable then

$$\frac{1}{2\pi}\text{Re}\left\{\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}(\Omega,\theta)d\omega\right\} = T(p_1^{-1})p_1^{\tau-1}\frac{1}{p_1 - p_2} + O(K_T^\tau \tau^{n_\alpha T}). \tag{4.319}$$

However, if p_2 is stable then

$$\frac{1}{2\pi}\text{Re}\left\{\int_I T^*(\Omega)e^{i\tau\omega}\widehat{H}(\Omega,\theta)d\omega\right\} = T(p_1^{-1})p_1^{\tau-1}\frac{1}{p_1 - p_2}$$

$$+ T(p_2^{-1})p_2^{\tau-1}\frac{1}{p_2 - p_1} + O(K_T^\tau \tau^{n_\alpha T})$$

$$= \frac{1}{p_1 - p_2}\left(T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1}\right) + O(K_T^\tau \tau^{n_\alpha T}). \tag{4.320}$$

For sufficiently large τ the signs of $T(p_1^{-1})$ and $T(p_2^{-1})$ do not change:

$$T(p_1^{-1}) = T(P_1^{-1}) + O(a_{1,\tau}),$$

$$T(p_2^{-1}) = T(P_2^{-1}) + O(a_{2,\tau}). \tag{4.321}$$

If the signs are the same then for all even τ, $P_1^{\tau-1} = -P_2^{\tau-1}$, or if the signs are different, then for all odd τ, $P_1^{\tau-1} = P_2^{\tau-1}$, and hence

$$\left| T(p_1^{-1})p_1^{\tau-1} - T(p_2^{-1})p_2^{\tau-1} \right| = \left| T(p_1^{-1})p_1^{\tau-1} \right| + \left| T(p_2^{-1})p_2^{\tau-1} \right| \tag{4.322}$$

(4.319) and (4.322) prove that inverting the unstable pole p_2 results asymptotically in a smaller cost function.

4.4.8.2   Case $p_1(\tau) = \dfrac{1}{p_2(\tau)}$

In this case the expression (4.69) is evaluated if p_2 is unstable. Then the result will be compared to the case when $\widehat{H}(z,\theta)$ has a double pole at the p_1. First, the unstable case is considered.

The assumption says that $p_2(\tau) = \frac{1}{p_1(\tau)}$. The term $T_3(\theta)$ can be written as (see (4.213))

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega,\theta)|^2 d\omega = \frac{p_1 + \dfrac{1}{p_1}}{\left(\dfrac{1}{p_1} - p_1 \right)^3}. \tag{4.323}$$

Substituting $p_1 = 1 - a_{1,\tau}$ the numerator and the denominator can be studied separately:

$$p_1 + \frac{1}{p_1} = 1 - a_{1,\tau} + \frac{1}{1 - a_{1,\tau}} = 2 + \frac{a_{1,\tau}^2}{1 - a_{1,\tau}} = 2 + O(a_\tau^2) \tag{4.324}$$

$$\left(\frac{1}{p_1} - p_1 \right)^3 = \left(\frac{a_{1,\tau}}{1 - a_{1,\tau}} + a_{1,\tau} \right)^3 = a_{1,\tau}^3 \left(\frac{1}{1 - a_{1,\tau}} + 1 \right)^3$$

$$= a_{1,\tau}^3 \left(2 + O(a_{1,\tau}) \right)^3 = a_{1,\tau}^3 (8 + O(a_{1,\tau})) \tag{4.325}$$

Hence the contribution of $T_3(\theta)$ in (4.69) is

$$\frac{2 + O(a_{1,\tau}^2)}{a_{1,\tau}^3(8 + O(a_{1,\tau}))} = \frac{1}{a_{1,\tau}^3} \left(\frac{1}{4} + O(a_{1,\tau}) \right) = \frac{1}{4a_{1,\tau}^3} + O\left(\frac{1}{a_{1,\tau}^2} \right) \tag{4.326}$$

The residue in $T_2(\theta,\tau)$ at p_1 can be calculated as

$$T(p_1^{-1})p_1^{\tau-1} \frac{1}{p_1 - \frac{1}{p_1}} = T(p_1^{-1})p_1^{\tau-1} \frac{-1}{a_{1,\tau}(2 + O(a_{1,\tau}))}. \tag{4.327}$$

Using the fact that $T(p_1^{-1}) = T(P_1^{-1}) + O(a_{1,\tau})$ the overall convergence rate of (4.69) is

$$\overline{C}(\theta, \tau) = \frac{\left(T(P_1^{-1})p_1^{\tau-1}\dfrac{1}{a_{1,\tau}}\dfrac{1}{2} + O(1) + O(K_T^\tau \tau^{n_\alpha T})\right)^2}{\dfrac{1}{4a_{1,\tau}^3} + O\left(\dfrac{1}{a_{1,\tau}^2}\right)} \tag{4.328}$$

Since $a_\tau = \frac{c_1}{\tau}$ and $O(1) + O(K_T^\tau \tau^{n_\alpha T}) = O(1)$ it leads to

$$\overline{C}(\theta, \tau) = (T(P_1^{-1})p_1^{\tau-1})^2 a_{1,\tau} + O(a_{1,\tau}^2) = (T(P_1^{-1}))^2 e^{-2c_1} c_1 \frac{1}{\tau} + O\left(\frac{1}{\tau^2}\right) \tag{4.329}$$

The first term in (4.329) should be maximized w.r.t. c_1:

$$\frac{\partial (T(P_1^{-1}))^2 e^{-2c_1} c_1}{\partial c_1} = (T(P_1^{-1}))^2(-2e^{-2c_1} c_1 + e^{-2c_1}) \tag{4.330}$$

$$-2e^{-2c_1} c_1 + e^{-2c_1} = 0. \tag{4.331}$$

The maximum can be achieved if $c_1 = \frac{1}{2}$. The maximal asymptotic value is

$$\frac{1}{\tau}(T(P_1^{-1}))^2 \frac{e^{-1}}{2} \tag{4.332}$$

Now we examine the case where $p_1(\tau) = p_2(\tau)$. The $\sum_{k=1}^{2} \text{Res}_{z=p_k} \widehat{H}(z, \theta) z^{-1} \widehat{H}(z^{-1}, \theta)$ can be written as (see (4.217))

$$\sum_{i=k}^{2} \text{Res}_{z=p_k} \widehat{H}(z, \theta) z^{-1} \widehat{H}(z^{-1}, \theta) = \frac{p_1^2 + 1}{(1 - p_1^2)^3}. \tag{4.333}$$

The numerator and the denominator are elaborated as

$$1 + p_1^2 = 1 + (1 - a_{1,\tau})^2 = 2 + O(a_{1,\tau}) \tag{4.334}$$

and

$$(1 - p_1^2) = 1 - (1 - a_{1,\tau})^2 = 2a_{1,\tau} - a_{1,\tau}^2 = a_{1,\tau}(2 + O(a_{1,\tau})). \tag{4.335}$$

In the term $T_2(\theta, \tau)$ the residue at the $z = p_1$ can be calculated by differentiation:

$$\left[T(z^{-1})z^{\tau-1}\right]' = \left[T\left(\frac{1}{z}\right)\right]' z^{\tau-1} + T\left(\frac{1}{z}\right)(\tau - 1)z^{\tau-2} \tag{4.336}$$

We know that

$$\left[T\left(\frac{1}{p_1}\right)\right]' p_1^{\tau-1} \tag{4.337}$$

is bounded because $T(z)$ and $T(z^{-1})$ are analytic in an appropriate neighborhood of $\mathbb{T}$. Since $O(K_T^\tau \tau^{n_\alpha T})$ is bounded as $\tau \to \infty$ the contribution of $T_2(\theta, \tau)$ is

$$= (\tau - 1)p_1^{\tau-2}T\left(\frac{1}{P_1}\right) + O(1) = \tau p_1^{\tau-2}T\left(\frac{1}{P_1}\right) + O(1). \tag{4.338}$$

Substituting $p_1 = 1 - a_{1,\tau}$, $a_{1,\tau} = \frac{c_1}{\tau}$, and using (4.69), $\overline{C}(\theta, \tau)$ can be expressed as

$$\overline{C}(\theta, \tau) = 4e^{-2c_1}\left(T(P_1^{-1})\right)^2 \frac{1}{\tau}c_1^3 + o\left(\frac{1}{\tau}\right). \tag{4.339}$$

$$\frac{\partial (T(P_1^{-1}))^2 4e^{-2c_1}c_1^3}{\partial c_1} = 4(T(P_1^{-1}))^2(-2e^{-2c_1}c_1^3 + 3c_1^2 e^{-2c_1}) \tag{4.340}$$

$$-2e^{-2c_1}c_1^3 + 3c_1^2 e^{-2c_1} = 0. \tag{4.341}$$

The maximum can be achieved if $c_1 = \frac{3}{2}$. The maximal asymptotic value is

$$\frac{1}{\tau}(T(P_1^{-1}))^2 \frac{27e^{-3}}{2} \tag{4.342}$$

Comparing (4.342) and (4.332) shows that the case of the stable double pole gives a higher cost function.

It means that the former case in which θ contains an unstable pole, cannot represent an optimal parameter vector θ_*.

4.4.8.3 Case $P_1 = P_2$ but $p_1(\tau) \neq \dfrac{1}{p_2(\tau)}$

This case is split again into two sub-cases:

- $a_{2,\tau}$ converges to the unit circle at the rate $\frac{1}{\tau}$, i.e. $a_{2,\tau} = \frac{c_2}{\tau}$,

- $a_{2,\tau}$ converges slower to the unit circle than $a_{1,\tau}$, hence $\frac{a_{1,\tau}}{a_{2,\tau}} = o(1)$.

In both cases it will be proved that asymptotically the cost function is the same. Moreover, it will be shown that this value cannot be optimal since in the case where $p_1(\tau) = p_2(\tau) = 1 - \frac{c_1}{\tau}$ asymptotically a lower cost function is achieved. The term $T_3(\theta)$ is the same as in (4.311). But in this case if $a_{2,\tau} = \frac{c_2}{\tau}$

$$p_1 + p_2 = P_1(1 - a_{1,\tau} + 1 + a_{2,\tau} + O(a_{2,\tau}^2))$$

$$= P_1\left(2 + \frac{-c_1 + c_2}{\tau} + O\left(\frac{1}{\tau^2}\right)\right) = 2P_1 + O\left(\frac{1}{\tau}\right) \tag{4.343}$$

and

$$p_1 - p_2 = P_1(1 - a_{1,\tau} - 1 - a_{2,\tau} + O(a_{2,\tau}^2)) =$$

$$= -P_1\left(\frac{c_1 + c_2}{\tau} + O\left(\frac{1}{\tau^2}\right)\right) = -P_1\frac{c_1 + c_2}{\tau}\left(1 + O\left(\frac{1}{\tau}\right)\right). \tag{4.344}$$

Collecting the results the term $T_3(\theta)$ (4.311) can be written as

$$= \tau^3\frac{2}{(c_1 + c_2)c_1c_24} + O(\tau^2) \tag{4.345}$$

The contribution of the term $T_2(\theta, \tau)$ at the $z = p_1$ is

$$T(p_1^{-1})p_1^{\tau-1}\frac{1}{p_1 - p_2} = T(P_1^{-1})e^{-c_1}\tau\frac{-1}{c_1 + c_2}\left(1 + O\left(\frac{1}{\tau}\right)\right). \tag{4.346}$$

Finally, $\overline{C}(\theta, \tau)$ can be expressed as

$$\frac{1}{\tau}(T(P_1^{-1}))^2e^{-2c_1}\frac{2c_1c_2}{c_1 + c_2} + O\left(\frac{1}{\tau^2}\right) + O(K_T^\tau \tau^{n_\alpha T}). \tag{4.347}$$

Since $c_1 > 0$, and $c_2 > 0$ the following upper bound is valid

$$\frac{c_1c_2}{c_1 + c_2} = \frac{c_1}{1 + \frac{c_1}{c_2}} \le c_1. \tag{4.348}$$

The upper bound of $e^{-2c_1}c_1$ is $\frac{1}{2}e^{-1}$. Therefore, using the fact that $O\left(\frac{1}{\tau^2}\right) + O(K_T^\tau \tau^{n_\alpha T}) = O\left(\frac{1}{\tau^2}\right)$ the asymptotic upper bound of $\overline{C}(\theta, \tau)$ can be written as

$$\frac{1}{\tau}(T(P_1^{-1}))^2e^{-1} + O\left(\frac{1}{\tau^2}\right). \tag{4.349}$$

Since $e^{-1} < \frac{27}{2}e^{-3}$, the previous case where $p_1 = p_2$ gives highest value of $\overline{C}(\theta, \tau)$. Therefore $\widehat{H}(\theta, z)$ which has a double pole at the point p_1 gives asymptotically lower cost function $C(\theta, \tau)$ (4.68).

Now we suppose that $a_{2,\tau}$ tends slower to 0 than $a_{1,\tau}$. The contribution of the factor $p_1 + p_2$ and $p_1 - p_2$ can be written as

$$p_1 + p_2 = 2P_1 + O(a_{2,\tau}) \tag{4.350}$$

and

$$p_1 - p_2 = -P_1a_{2,\tau}(1 + o(1)). \tag{4.351}$$

The overall contribution of the term T_3 is

$$= \frac{2P_1 + O(a_{2,\tau})}{-P_1 a_{2,\tau}(1 + o(1))P_1 a_{2,\tau}(-2 + O(a_{2,\tau}))\frac{c_1}{\tau}P_1\left(2 + O\left(\frac{1}{\tau}\right)\right)}$$

$$= \frac{\tau}{2c_1 a_{2,\tau}^2} + o\left(\frac{\tau}{a_{2,\tau}^2}\right). \tag{4.352}$$

In the term $T_2(\theta, \tau)$, the residue at p_1 yields

$$T(p_1^{-1})p_1^{\tau-1}\frac{1}{p_1 - p_2} = T(P_1^{-1})e^{-c_1}(1 + o(1))\frac{1}{a_{2,\tau}(1 + o(1))}$$

$$= \frac{T(P_1^{-1})e^{-c_1}}{a_{2,\tau}} + o\left(\frac{1}{a_{2,\tau}}\right). \tag{4.353}$$

So $\overline{C}(\theta, \tau)$ can be written as

$$\overline{C}(\theta, \tau) = \frac{2c_1\left(T(P_1^{-1})e^{-c_1}\right)^2}{\tau} + o\left(\frac{1}{\tau}\right) + O\left(\frac{a_{2,\tau}^2}{\tau}K_T^\tau \tau^{n_{\alpha T}}\right)$$

$$= \frac{(T(P_1^{-1}))^2}{\tau}2c_1 e^{-2c_1} + o\left(\frac{1}{\tau}\right). \tag{4.354}$$

Hence, the same argument like in the previous case can be applied.

4.4.8.4 Case $p_2(\tau)$ does not approach $\mathbb{T}$

In this step the proof starts again from (4.311). We assume that $p_2(\tau) \to Q$ where $|Q| > 1$. It will turn out that the result and, hence, the conclusion of this subsection are the same like in the previous case. Now

$$p_1 + p_2 = P_1 + Q + O\left(\frac{1}{\tau} + a_{2,\tau}\right),$$

$$p_1 - p_2 = P_1 - Q + O\left(\frac{1}{\tau} + a_{2,\tau}\right),$$

$$1 - p_2^2 = 1 - Q^2 + O(a_{2,\tau}). \tag{4.355}$$

Collecting the results the term $T_3(\theta)$ (4.311) can be written as

$$\frac{\left(P_1 + Q + O\left(\frac{1}{\tau} + a_{2,\tau}\right)\right)}{\left(P_1 - Q + O\left(\frac{1}{\tau} + a_{2,\tau}\right)\right)\frac{c_1}{\tau}\left(2 + O\left(\frac{1}{\tau}\right)\right)\left(1 - Q^2 + O(a_{2,\tau})\right)}$$

$$= \frac{\tau}{2c_1}\frac{(P_1 + Q)}{(P_1 - Q)(1 - Q^2)} + O\left(\frac{1}{\tau} + a_{2,\tau}\right). \tag{4.356}$$

The residue of $T_2(\theta, \tau)$ at the $z = p_1$ is

$$T(p_1^{-1})p_1^{\tau-1}\frac{1}{p_1 - p_2} = T(P_1^{-1})e^{-c_1}\frac{1}{P_1 - Q} + o(1). \tag{4.357}$$

Finally, the cost function can be expressed as

$$\frac{2c_1}{\tau}(T(P_1^{-1}))^2 e^{-2c_1}\frac{1 - Q^2}{(P_1 - Q)(P_1 + Q)}$$

$$+ o\left(\frac{1}{\tau}\right) + O\left(K_T^\tau \tau^{n_\alpha T}\right) + O\left(\frac{a_{2,\tau}^2}{\tau}K_T^\tau \tau^{n_\alpha T}\right). \tag{4.358}$$

Using the facts that $(P_1 - Q)(P_1 + Q) = P_1^2 - Q^2 = 1 - Q^2$ and

$$o\left(\frac{1}{\tau}\right) + O\left(K_T^\tau \tau^{n_\alpha T}\right) + O\left(\frac{a_{2,\tau}^2}{\tau}K_T^\tau \tau^{n_\alpha T}\right) = o\left(\frac{1}{\tau}\right) \tag{4.359}$$

(4.358) becomes

$$= \frac{2c_1}{\tau}(T(P_1^{-1}))^2 e^{-2c_1} + o\left(\frac{1}{\tau}\right). \tag{4.360}$$

Therefore, the same argument as in the previous cases can be applied.

4.5   Generalization

In the proof presented in the previous section, the cases of different orders are treated separately. This method obviously does not work for further n_β's and n_α's. For the general case the same approach cannot be applied because of the increasing complexity of higher orders. Instead, the general proof will be carried out for a somewhat restricted set of cases. The results presented in this section are based on the following assumptions:

Assumption 4.5.1. *If a pole $p(\tau)$ of $D(\Omega, \theta_*)$ converges to the point P on $\mathbb{T}$ as $\tau \to \infty$ then $\theta_*(\tau)$ does not define another pole $q(\tau)$ such that the limit set of the sequence $q(\tau)$ contains P.*

Assumption 4.5.2. *Every pole q of $D(\Omega, \theta_*)$ converges to $\mathbb{T}$ so that their contributions $T(z^{-1})z^{\tau-1}H(z,\theta)\big|_{z=q}$ in the numerator have the same convergence rate.*

In this section we will show theorem 4.3.1 for arbitrary n_β and n_α under the above assumptions. As illustrations of the reasonability of the assumptions, many examples are shown in the thesis [Vuerinckx, 1998], which confirms the asymptotic behavior of the poles. In figure 4.3 such an example is presented. The optimization is based on methods to be described in the next chapter. The specifications of this particular example are: the frequency band is $[0.05 : 0.005 : 0.5]$ and the target transfer function is $T(\omega) = 1$. In the given band the transfer function was approximated by a 10th-order model with $\tau = 30$.
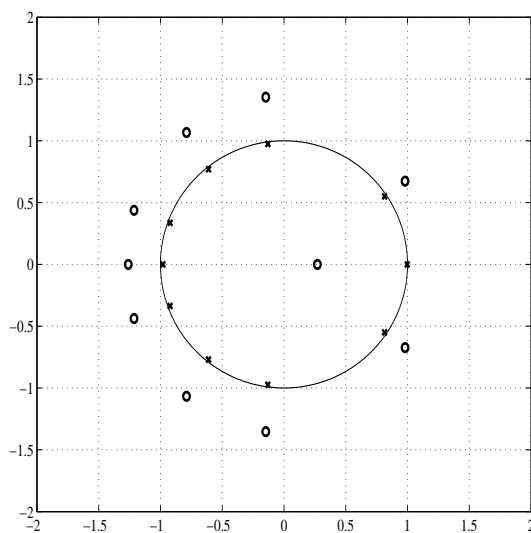
Figure 4.3: Demonstration of asymptotic behavior of the poles. Marks 'x' denote the poles, marks 'o' denote the zeros.

In the previous sections the following statements have been proved without using the assumption $n_\beta + n_\alpha \leq 2$, thus here we will only refer to them:

- $T_2(\theta, \tau)$ and $T_3(\theta)$ are bounded.

- The zeros and poles of the optimal parameter vector sequence $\theta_*(\tau)$, which converge to infinity, can be disregarded.

- There exists at least one pole converging to the unit circle.

- We can construct a $\theta(\tau)$ such that $\overline{C}(\theta, \tau) \sim \frac{1}{\tau}$. Hence, for an optimal parameter vector sequence $\theta_*(\tau)$, $\tau\overline{C}(\theta_*, \tau)$ does not converge to 0.

These statements are valid for the general case. The remaining steps presented here are very similar to those in the proof for the case $n_\beta + n_\alpha \leq 2$. Roughly, the following statements will be discussed and proved:

- The maximal convergence rate of an optimal parameter vector sequence is $O\left(\frac{1}{\tau}\right)$.

- Inverting an unstable subsequence of a pole converging to $\mathbb{T}$ does not change the asymptotic value of $T_3(\theta)$.

- Inverting an unstable subsequence of a pole converging to $\mathbb{T}$ increases the asymptotic value of the numerator of $\overline{C}(\theta, \tau)$.

The last two steps can be applied repeatedly. Then, assuming that θ_* defines an unstable pole sequence $p_k(\tau)$, a new $\tilde{p}_k(\tau)$ can be constructed such that $\overline{C}(\tilde{\theta}, \tau) > \overline{C}(\theta_*, \tau)$, i.e. $C(\tilde{\theta}, \tau) < C(\theta_*, \tau)$. This is a contradiction with global minimization of $C(\theta_*, \tau)$, therefore we can conclude that an optimal parameter vector sequence $\theta_*(\tau)$ does not define any unstable pole subsequence.

4.5.1   Maximum Convergence Rate

As mentioned above, in the previous section a construction of parameter vector sequence $\theta(\tau)$ was shown such that $\overline{C}(\theta, \tau) \sim \frac{1}{\tau}$. Hence, it is also an upper bound for all the optimal parameter vector sequence $\theta_*(\tau)$. This part of the thesis is devoted to prove that this bound is exact, i.e. for every parameter vector sequence $\theta_*(\tau)$ which minimize the cost function, $\overline{C}(\theta_*, \tau) \sim \frac{1}{\tau}$. A practical consequence is that an optimal parameter vector sequence $\theta_*(\tau)$ cannot determine a pole sequence which converges to the unit circle at rate faster than $\frac{1}{\tau}$.

The statements about the maximal convergence rate can be proved without the assumption 4.5.2. Therefore, during the steps of the proof of maximal convergence rate following subsets of poles have to be distinguished. Let us assume that the first N poles converge to $\mathbb{T}$ and this subset of $\mathcal{P}$ is denoted by $\mathcal{P}_{\mathbb{T}}$. $\mathcal{P}$ can be written as disjoint union of the sets $\mathcal{P}_s$ and $\mathcal{P}_u$, i.e. $\mathcal{P} = \mathcal{P}_s \cup \mathcal{P}_u$ with $\mathcal{P}_s \cap \mathcal{P}_u = \emptyset$, where $\mathcal{P}_s$ contains the poles that are inside the unit circle and $\mathcal{P}_u$ contains the unstable poles. Moreover, without losing generality, we can assume that all the sequences $p_k(\tau)$, $k = 1, \ldots, n_\alpha$ and $z_l(\tau)$, $l = 1, \ldots, n_\beta$, constitute convergent sequences. Their limits are denoted P_k, $k = 1, \ldots, n_\alpha$ for poles, and Z_l, $l = 1, \ldots, n_\beta$ for zeros. It is worth noting that for all $p_k \in \mathcal{P}_{\mathbb{T}}$, $|P_k| = 1$ because $P_k \in \mathbb{T}$. Like in the previous cases, we assume that the zeros and poles can be written as

$$z_k = Z_k(1 - b_{k,\tau}) \text{ and } p_k = P_k(1 - a_{k,\tau}). \tag{4.361}$$

4.5.1.1   Investigation of the Numerator of $\overline{C}(\theta, \tau)$

The numerator of $\overline{C}(\theta, \tau)$ (4.69) can be written as

$$\int_I T^*(\Omega) e^{i\tau\omega} \widehat{H}(\Omega, \theta) d\omega = 2\pi \left(\sum_{k, p_k \in (\mathcal{P}_s \cap \mathcal{P}_{\mathbb{T}})} \operatorname{Res}_{z=p_k} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) \right.$$

$$\left. + \sum_{k, p_k \in (\mathcal{P}_s \setminus \mathcal{P}_{\mathbb{T}})} \operatorname{Res}_{z=p_k} T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta) \right) \tag{4.362}$$

where $\widehat{H}(z,\theta)$ is defined in (4.63). Since $T(z^{-1})$ does not change as $\tau \to \infty$, it has no pole converging to $\mathbb{T}$ or to ∞. It means that for all $p_k \in \mathcal{P}_s$, and for all $l = 0, \ldots, n_\alpha - 1$, the expression

$$\left[T(z^{-1})\right]^{(l)}\bigg|_{z=p_k} \tag{4.363}$$

is bounded. If $p_k \in \mathcal{P}_s\backslash\mathcal{P}_\mathbb{T}$ then applying note 4.2.8 to $p_k^{\tau-l-1}$ for $l = 0, \ldots, n_\alpha - 1$ shows that

$$\left[z^{\tau-1}\right]^{(l)}\bigg|_{z=p_k} = (\tau-1)\cdots(\tau-l)p_k^{\tau-l-1} \tag{4.364}$$

converges to 0 at least at exponential rate.

Using the results of §4.4.4 we know that $\forall k, p_k \in \mathcal{P}\backslash\mathcal{P}_\mathbb{T}$ and for all $l = 1, \ldots, n_\alpha$

$$\left[\widehat{H}(z,\theta)(z-p_k)^l\right]^{(l-1)}\bigg|_{z=p_k} \tag{4.365}$$

is lower bounded. The upper bound is

$$\frac{\left(\displaystyle\prod_{m=1}^{n_\beta} \max\left(1, P_k - Z_m\right)\right)^l}{\left(\displaystyle\prod_{m\neq k, p_m \in \mathcal{P}} \min\left(1, P_k - P_m\right)\right)^l}. \tag{4.366}$$

for all τ sufficiently large.

So, if $p_k \in \mathcal{P}\backslash\mathcal{P}_\mathbb{T}$ then

$$\left|\mathrm{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right| \tag{4.367}$$

converges to 0 at least at exponential rate.

If $p_k \in \mathcal{P}_\mathbb{T}$, then, according to notes 4.2.6 and 4.2.7, the factor $|p_k|^{\tau-1}$ converges to a non-zero finite value. Since $\left|\frac{1}{P_k}\right| = 1$ and $T(z^{-1})$ has no pole on $\mathbb{T}$, $T(z^{-1})$ is analytic in the neighborhood of $\frac{1}{P_k}$. Therefore,

$$T\left(\frac{1}{p_k}\right) \to T\left(\frac{1}{P_k}\right), \text{ i.e. } T\left(\frac{1}{p_k}\right) = T\left(\frac{1}{P_k}\right)(1 + o(1)). \tag{4.368}$$

Using the assumption 4.5.1, we have that

$$\left[T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)(z-p_k)\right]\bigg|_{z=p_k} = \frac{N(p_k,\theta)p_k^{\tau-1}}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha}(p_k - p_l)}T\left(\frac{1}{p_k}\right)$$

$$
= \frac{\displaystyle\prod_{l,Z_l=P_k} (-a_{k,\tau} + b_{l,\tau})p_k^{\tau-1} \prod_{l,Z_l\neq P_k} (P_k - Z_l)(1 + o(1))}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha} (P_k - P_l)} T\left(\frac{1}{P_k}\right) \tag{4.369}
$$

If $p_k(\tau)$ is a real sequence, then

$$
p_k^{\tau-1} = (1 - a_{k,\tau})^{\tau-1} \tag{4.370}
$$

and

$$
\frac{\displaystyle\prod_{l,Z_l\neq P_k} (P_k - Z_l)}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha} (P_k - P_l)} T\left(\frac{1}{P_K}\right) \tag{4.371}
$$

is a finite and real number.

If it $p_k(\tau)$ is a complex sequence then

$$
p_k^{\tau-1} = (1 - |a_{k,\tau}|)^{\tau-1} e^{i\phi_{k,\tau}(\tau-1)} \tag{4.372}
$$

with $\phi_{k,\tau} = \arg(p_k(\tau))$, and

$$
\left[\widehat{H}(z,\theta)T\left(\frac{1}{z}\right) z^{\tau-1}(z - p_k)\right]\Bigg|_{z=p_k} + \left[\widehat{H}(z,\theta)z^{\tau-1}T\left(\frac{1}{z}\right)(z - p_k^*)\right]\Bigg|_{z=p_k^*}
$$

$$
= \mathrm{Re}\left\{ \frac{\displaystyle\prod_{l,Z_l=P_k} (-a_{k,\tau} + b_{l,\tau})e^{i\phi_{k,\tau}(\tau-1)}}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha} (P_k - P_l)} \prod_{l,Z_l\neq P_k} (P_k - Z_l)T\left(\frac{1}{P_k}\right) \right\}
$$

$$
\times 2(1 - |a_{k,\tau}|)^{\tau-1}(1 + o(1)). \tag{4.373}
$$

Applying the usual estimation $|\mathrm{Re}\{z\}| \leq |z|$, we have

$$
\left| \left[\widehat{H}(z,\theta)T\left(\frac{1}{z}\right) z^{\tau-1}(z - p_k)\right]\Bigg|_{z=p_k} + \left[\widehat{H}(z,\theta)T\left(\frac{1}{z}\right) z^{\tau-1}(z - p_k^*)\right]\Bigg|_{z=p_k^*} \right|
$$

$$
\leq (1 - |a_{k,\tau}|)^{\tau-1} 2 \frac{\displaystyle\prod_{l,Z_l=P_k} (|a_{k,\tau}| + |b_{l,\tau}|) \prod_{l,Z_l\neq P_k} |P_k - Z_l|}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha} |P_k - P_l|} \left|T\left(\frac{1}{P_k}\right)\right| (1 + o(1)). \tag{4.374}
$$

The overall estimation of the numerator of $\overline{C}(\theta,\tau)$ is bounded as

$$\left|\sum_{k=1}^{n_\alpha}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|^2$$

$$\leq\left|\sum_{k,p_k\in\mathcal{P}_\mathbb{T}}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|^2+\left|\sum_{k,p_k\in\mathcal{P}\backslash\mathcal{P}_\mathbb{T}}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|^2+$$

$$+2\left|\sum_{k,p_k\in\mathcal{P}_\mathbb{T}}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|\left|\sum_{k,p_k\in\mathcal{P}\backslash\mathcal{P}_\mathbb{T}}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right| \quad (4.375)$$

The second and the third terms converge to 0 at least at exponential rate. The first one can be upper bounded as

$$\left|\sum_{k,p_k\in\mathcal{P}_\mathbb{T}}\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|^2$$

$$\leq\sum_{k,p_k\in\mathcal{P}_\mathbb{T}}\sum_{m,p_m\in\mathcal{P}_\mathbb{T}}\left|\operatorname{Res}_{z=p_k}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|\left|\operatorname{Res}_{z=p_m}T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)\right|$$

$$\leq\sum_{k,p_k\in\mathcal{P}_\mathbb{T}}\sum_{m,p_m\in\mathcal{P}_\mathbb{T}}4(1-|a_{k,\tau}|)^{\tau-1}(1-|a_{m,\tau}|)^{\tau-1}\prod_{l,Z_l=P_k}(|a_{k,\tau}|+|b_{l,\tau}|)$$

$$\times\prod_{l,Z_l=P_m}(|a_{m,\tau}|+|b_{l,\tau}|)\frac{\displaystyle\prod_{l,Z_l\neq P_k}|P_k-Z_l|\prod_{l,Z_l\neq P_m}|P_m-Z_l|}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha}|P_k-P_l|\prod_{l=1,l\neq m}^{n_\alpha}|P_m-P_l|}(1+o(1)) \quad (4.376)$$

The upper bound can be used to determine the convergence rate of the numerator of $\overline{C}(\theta,\tau)$. The convergence rate is determined by the factors $(1-|a_{k,\tau}|)^{\tau-1}$, $(1-|a_{m,\tau}|)^{\tau-1}$, $(|a_{k,\tau}|+|b_{l,\tau}|)$ and $(|a_{m,\tau}|+|b_{l,\tau}|)$. It is worth noting that if at least in one factor $(1-|a_{k,\tau}|)^{\tau-1}$ the convergence rate of $a_{k,\tau}$ is slower than $\frac{1}{\tau}$, then according to note 4.2.8 the overall convergence rate is exponential. Otherwise, the factors $(1-|a_{k,\tau}|)^{\tau-1}$ converge to a positive number, and the overall convergence rate is determined by product of the factors $(|a_{k,\tau}|+|b_{l,\tau}|)$.

4.5.1.2   Investigation of the Denominator of $\overline{C}(\theta, \tau)$

The dominator of $\overline{C}(\theta, \tau)$ equals to $\| \widehat{H}(\Omega, \theta) \|_2^2 = \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$. As we seen in (4.52) the integral can be written as sum of residues calculated at the poles inside the unit circle:

$$\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = 2\pi \sum_{k, p_k \in \mathbb{P}_D} \text{Res}_{z=p_k} \widehat{H}\left(z^{-1}, \theta\right) z^{-1} \widehat{H}(z, \theta). \tag{4.377}$$

In this section the asymptotic properties of $\| \widehat{H}(\Omega, \theta) \|_2^2$ are studied. We need to consider two different cases:

- $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$ remains bounded as $\tau \to \infty$, or

- $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega \to \infty$ as $\tau \to \infty$.

In both cases our aim is calculating an appropriate lower-bound. The trivial lower bound is 0, but it cannot be used to determine a useful upper-bound of $\overline{C}(\theta, \tau)$.

In the first cases, a short calculation shows that $\| \widehat{H}(\Omega, \theta) \|_2^2 \geq C_1 > 0$ where C_1 is a constant depends on the asymptotic poles and zeros configuration. Th calculation was described in §4.4.4.1. The result is

$$\int_I \left|\widehat{H}(e^{i\omega}, \theta)\right|^2 d\omega \geq \frac{1}{(M+1)^{2n_\alpha}} K_Z^2 \frac{2\pi}{2n_\beta + 1}. \tag{4.378}$$

Since $\widehat{H}(\Omega, \theta)$ has only bounded poles and zeros, in the second case, the limit value shows that at least one pole converges to $\mathbb{T}$. The presence of the poles $k = 1, \ldots, N$ implies that $T_3(\theta)$ converges to infinity. Now, we will show that the contributions of these kinds of poles can be summarized and the convergence rate of a contribution depends only on the pole at which the residue are calculated. It is worth noting that $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$ is always positive, hence if it converges to infinity then there exists at least one pole or a complex pole pair such that

$$\text{Res}_{z=p_1} \widehat{H}(z, \theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z}, \theta\right) \to \infty \tag{4.379}$$

or

$$\text{Res}_{z=p_1} \widehat{H}(z, \theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z}, \theta\right) + \text{Res}_{z=p_2} \widehat{H}(z, \theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z}, \theta\right) \to \infty, \tag{4.380}$$

where $p_2 = p_1^*$.

As we already seen $\| \widehat{H}(\Omega, \theta) \|_2^2 \to \infty$ is possible only if at least one pole converges to the unit circle. Residues calculated at the poles which converge to $\mathbb{T}$ can converge to infinity. Because of assumption 4.5.1 for all τ sufficiently large the poles converging to $\mathbb{T}$

are single. Using (4.5) residue at p_1, which converge to $\mathbb{T}$, can be written as

$$\text{Res}_{z=p_1} \widehat{H}(z,\theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z},\theta\right) = \frac{N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)}{\left(\frac{1}{p_1} - p_1\right) p_1 \displaystyle\prod_{p_k, p_k \neq p_1} (p_1 - p_k)\left(\frac{1}{p_1} - p_k\right)}, \quad (4.381)$$

or

$$\text{Res}_{z=p_1} \widehat{H}(z,\theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z},\theta\right)$$

$$= \frac{N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)}{\left(\frac{1}{p_1} - p_1^*\right) p_1(p_1 - p_1^*)\left(\frac{1}{p_1} - p_1\right) \displaystyle\prod_{p_k, p_k \neq p_1, p_k \neq p_1^*} (p_1 - p_k)\left(\frac{1}{p_1} - p_k\right)}, \quad (4.382)$$

In the real case the factor $\left(\frac{1}{p_1} - p_1\right)$ converges to infinity. In the case of complex conjugate pole pairs $\left(\frac{1}{p_1} - p_1^*\right)$ converges to infinity.

In the numerator $N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)$ it is possible to be zeros which converge to the same point like p_1. Therefore, it is examined separately. After this, the remaining factors in (4.381) and in (4.382) are studied.

4.5.1.3 Analysis of $N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)$

First, it is noted that the factor $N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)$ is invariant when replacing p_1 by $\frac{1}{p_1}$.

In the case of a complex conjugate pole pair we are using the fact that the numerator polynomial $N(z)$ has real coefficients, hence $N(p_1^*) = N^*(p_1)$. Since

$$p_k \to P_k \quad \Rightarrow \quad \frac{1}{p_k} \to P_k^*, \quad (4.383)$$

we have

$$N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right) = \prod_{l, Z_l = P_1} Z_l\left(b_{l,\tau} - a_{1,\tau}\right) \prod_{l, Z_l \neq P_1} (P_1 - Z_l + O(a_{1,\tau} - b_{l,\tau}))$$

$$\prod_{l, Z_l^* = P_1} Z_l^*\left(\frac{1}{1-a_{1,\tau}} - (1 - b_{l,\tau})\right) \prod_{l, Z_l^* \neq P_1} \left(\frac{1}{P_1(1-a_{1,\tau})} - Z_l^*(1 - b_{l,\tau})\right). \quad (4.384)$$

The factor $\left(\frac{1}{1-a_{1,\tau}} - (1 - b_{l,\tau}) \right)$ can be written as

$$\frac{1}{1 - a_{1,\tau}} - (1 - b_{l,\tau}) = 1 + a_{1,\tau} + O(a_{1,\tau}^2) - (1 - b_{l,\tau}) = a_{1,\tau} + b_{l,\tau} + O(a_{1,\tau}^2). \quad (4.385)$$

The set of zeros can be divided into the following disjoint subsets:

- $\mathcal{Z}_{o,k}$: $z_l \in \mathcal{Z}_{o,k}$ if and only if $Z_l \neq P_k$. This subset contains the zeros that do not converge to P_k.

- $\mathcal{Z}_{s,k}$: $z_l \in \mathcal{Z}_{s,k}$ if and only if $Z_l = P_k$ and $a_{k,\tau} = o(b_{l,\tau})$. The zeros belonging to this set converge to P_k but at a slower rate than p_k does.

- $\mathcal{Z}_{e,k}$: $z_l \in \mathcal{Z}_{e,k}$ if and only if $Z_l = P_k$ and $b_{l,\tau} \sim a_{k,\tau}$, i.e. the corresponding zeros converge to P_k at exactly the same rate as p_k does.

- $\mathcal{Z}_{f,k}$: $z_l \in \mathcal{Z}_{f,k}$ if and only if $Z_l = P_k$ and $b_{l,\tau} = o(a_{k,\tau})$. In this subset are the zeros which converge to P_k faster than p_k does.

The convergence rate defines an equivalence relation and this ensures that the defined subsets are disjoint.

Supposing that $z_l \in \mathcal{Z}_{s,k}$ we have

$$b_{l,\tau} \pm a_{1,\tau} = b_{l,\tau}(1 + o(1)), \quad (4.386)$$

and if $z_l \in \mathcal{Z}_{o,k}$

$$\frac{1}{P_1(1 - a_{1,\tau})} - Z_l^*(1 - b_{l,\tau}) = P_1^* - Z_l^* + O(a_{1,\tau} + b_{1,\tau}). \quad (4.387)$$

In the case where $z_l \in \mathcal{Z}_{f,k}$,

$$b_{l,\tau} \pm a_{1,\tau} = \pm a_{1,\tau}(1 + o(1)). \quad (4.388)$$

If $z_l \in \mathcal{Z}_{e,k}$, then it is possible that

$$b_{l,\tau} \pm a_{1,\tau} = o(a_{1,\tau}) = o(b_{l,\tau}), \quad (4.389)$$

and only the following upper bound can be constructed

$$(p_1 - z_l)\left(p_1 - \frac{1}{z_l} \right) = (b_{l,\tau} - a_{1,\tau})(-b_{l,\tau} - a_{1,\tau})(1 + o(1)) = O(a_{1,\tau}^2). \quad (4.390)$$

We conclude that (4.384) can be approximated as

$$= \prod_{l,z_l \in \mathcal{Z}_{s,k}} b_{l,\tau}^2 \prod_{l,z_l \in \mathcal{Z}_{e,k}} (b_{l,\tau} - a_{k,\tau})(-b_{l,\tau} - a_{k,\tau})$$

$$\times \prod_{l,z_l \in \mathcal{Z}_{f,k}} a_{k,\tau}^2 \prod_{l,z_l \in \mathcal{Z}_{o,k}} |P_k - Z_l|^2 (1 + o(1)). \tag{4.391}$$

It is worth noting that if at least one of the sets $\mathcal{Z}_{e,k}$ and $\mathcal{Z}_{f,k}$ is not empty, then

$$N(p_k, \theta) N\left(\frac{1}{p_k}, \theta\right) = o(a_{k,\tau}). \tag{4.392}$$

4.5.1.4   Real Pole

In the case of a real pole $p_1 = P_1(1 - a_{1,\tau})$ with $|P_1| = 1$,

$$\text{Res}_{z=p_1} \widehat{H}(z, \theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z}, \theta\right) = \frac{N(p_1) N\left(\dfrac{1}{p_1}\right)}{p_1 \left(\dfrac{1}{p_1} - p_1\right) \prod_{k=2}^{n_\alpha} (p_1 - p_k)\left(\dfrac{1}{p_1} - p_k\right)}. \tag{4.393}$$

Since P_1 is real, and $|P_1| = 1$, it follows that $P_1 = 1$ or $P_1 = -1$. Assumption 4.5.1 implies that for $k = 2, \ldots, n_\alpha$:

$$(p_1 - p_k) = P_1 - P_k + o(1) \tag{4.394}$$

and

$$\left(\frac{1}{p_1} - p_k\right) = P_1 - P_k + o(1). \tag{4.395}$$

If P_k is complex then there exists an l, $2 \le l \le n_\alpha$, $l \ne k$ such that $P_l = P_k^*$. Hence

$$(p_1 - p_k)(p_1 - p_l) = (P_1 - P_k + o(1))(P_1 - P_k^* + o(1)) = |P_1 - P_k|^2 + o(1) \tag{4.396}$$

and

$$\left(\frac{1}{p_1} - p_k\right)\left(\frac{1}{p_1} - p_l\right) = (P_1 - P_k + o(1))(P_1 - P_k^* + o(1)) = |P_1 - P_k|^2 + o(1). \tag{4.397}$$

The first factor in the denominator of (4.393)

$$p_1\left(\frac{1}{p_1} - p_1\right) = 1 - p_1^2 = 1 - P_1^2(1 - a_{1,\tau})^2 \tag{4.398}$$

$$= 2a_{1,\tau} + O(a_{1,\tau}^2) = 2a_{1,\tau}(1 + o(1)) = 2\text{Re}\{a_{1,\tau}\}(1 + o(1)) \tag{4.399}$$

where in the last step we used the assumption that $a_{1,\tau}$ is real.

Therefore the overall contribution of the residue calculated at p_1 is

$$\text{Res}_{z=p_1}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) = \frac{N(p_1,\theta)N\left(\dfrac{1}{p_1},\theta\right)}{2\text{Re}\{a_{1,\tau}\}}\frac{1}{\prod_{k=2}^{n_\alpha}(|P_1-P_k|^2)}(1+o(1)). \quad (4.400)$$

The convergence rate of the contribution is determined by the factor

$$\frac{N(p_1,\theta)N\left(\dfrac{1}{p_1},\theta\right)}{2\text{Re}\{a_{1,\tau}\}}. \quad (4.401)$$

4.5.1.5   Complex Conjugate Pole Pair

In the case of a complex conjugate pole pair, the sum of the following terms have to be studied:

$$\text{Res}_{z=p_1}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) + \text{Res}_{z=p_2}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right). \quad (4.402)$$

Evaluating the residues,

$$= \frac{N(p_1,\theta)N\left(\dfrac{1}{p_1},\theta\right)}{p_1\left(\dfrac{1}{p_1}-p_1\right)(p_1-p_1^*)\left(\dfrac{1}{p_1}-p_1^*\right)\displaystyle\prod_{k=3}^{n_\alpha}(p_1-p_k)\left(\dfrac{1}{p_1}-p_k\right)}$$

$$+ \frac{N(p_1^*,\theta)N\left(\dfrac{1}{p_1^*},\theta\right)}{p_1^*\left(\dfrac{1}{p_1^*}-p_1^*\right)(p_1^*-p_1)\left(\dfrac{1}{p_1^*}-p_1\right)\displaystyle\prod_{k=3}^{n_\alpha}(p_1^*-p_k)\left(\dfrac{1}{p_1^*}-p_k\right)}. \quad (4.403)$$

We will prove that the residue at p_1 is a real number and therefore equals to the residue calculated at p_1^*.

In the denominator we can use the fact that for all factors $(p_k - p_l)$, there exists a factor in the product which equals to $\left(\frac{1}{p_k}-p_l^*\right)$. Therefore,

$$\prod_{k=3}^{n_\alpha}(p_1-p_k)\left(\frac{1}{p_1}-p_k\right) = \prod_{k,p_k\in\mathcal{P}_{k,\mathbb{R}}}(p_1-p_k)\left(\frac{1}{p_1}-p_k\right)$$

$$\times \prod_{k,p_k\in\mathcal{P}_{k,\mathbb{C}}}(p_1-p_k)\left(\frac{1}{p_1}-p_k^*\right)(p_1-p_k^*)\left(\frac{1}{p_1}-p_k\right) \quad (4.404)$$

where $\mathcal{P}_{k,\mathbb{R}}$ and $\mathcal{P}_{k,\mathbb{C}}$ are subsets of $\mathcal{P}$ and are defined as

$$p_k \in \mathcal{P}_{k,\mathbb{R}} \text{ if } \lim_{\tau \to \infty} p_k = P_k \text{ and } \forall \tau : \text{Im}\{p_k\} = 0, \tag{4.405}$$

$$p_k \in \mathcal{P}_{k,\mathbb{C}} \text{ if } \lim_{\tau \to \infty} p_k = P_k \text{ and } \forall \tau : \text{Im}\{p_k\} \neq 0. \tag{4.406}$$

It is worth noting that in both cases we need not assume that the poles are stable.

In the case of $p_k \in \mathcal{P}_{k,\mathbb{R}}$

$$(p_1 - p_k)\left(\frac{1}{p_1} - p_k\right) = (P_1 - P_k + O(-a_{1,\tau} + a_{k,\tau}))(P_1^* - P_k + O(a_{1,\tau} + a_{k,\tau}))$$

$$= |P_1 - P_k|^2(1 + o(1)) \tag{4.407}$$

and similarly for $k, p_k \in \mathcal{P}_{k,\mathbb{C}}$ we have

$$(p_1 - p_k)\left(\frac{1}{p_1} - p_k^*\right) = |P_1 - P_k|^2(1 + o(1)) \tag{4.408}$$

and

$$(p_1 - p_k^*)\left(\frac{1}{p_1} - p_k\right) = |P_1^* - P_k|^2(1 + o(1)). \tag{4.409}$$

By collecting the results,

$$\frac{N(p_1, \theta)N\left(\frac{1}{p_1}, \theta\right)}{\prod\limits_{k=3}^{n_\alpha}(p_1 - p_k)\left(\frac{1}{p_1} - p_k\right)}$$

$$= \frac{N(p_1, \theta)N\left(\frac{1}{p_1}, \theta\right)}{\prod\limits_{k,p_k \in \mathcal{P}_{k,\mathbb{R}}}|P_1 - P_k|^2 \prod\limits_{k,p_k \in \mathcal{P}_{k,\mathbb{C}}}|P_1 - P_k|^2|P_1^* - P_k|^2}(1 + o(1)). \tag{4.410}$$

It is worth noting that the last expression is positive for all τ sufficiently large and that if we replace p_1 by p_1^* then the last expression remains the same. The asymptotic behavior of $\frac{1}{p_1}$ is

$$\frac{1}{p_1} = \frac{1}{P_1(1 - a_{1,\tau})} = P_1^*\left(1 + a_{1,\tau} + \frac{a_{1,\tau}^2}{1 - a_{1,\tau}}\right) = P_1^*\left(1 + a_{1,\tau} + O(a_{1,\tau}^2)\right), \tag{4.411}$$

and substituting the results into $\left(\frac{1}{p_1} - p_1^*\right)$ leads to

$$\left(\frac{1}{p_1} - p_1^*\right) = P_1^* \left(1 + a_{1,\tau} + O(a_{1,\tau}^2)\right) - P_1^*(1 - a_{1,\tau}^*) = P_1^* 2\mathrm{Re}\{a_{1,\tau}\}\left(1 + o(1)\right). (4.412)$$

The first four factors in the denominator of the first term in (4.403) can be written as

$$p_1 \left(\frac{1}{p_1} - p_1\right)(p_1 - p_1^*)\left(\frac{1}{p_1} - p_1^*\right)$$

$$= P_1(1 + o(1))(P_1^* - P_1 + O(a_{1,\tau}))(P_1 - P_1^* + O(a_{1,\tau}))P_1^*(2\mathrm{Re}\{a_{1,\tau}\} + o(1))$$

$$= |P_1|^2|P_1 - P_1^*|^2 2\mathrm{Re}\{a_{1,\tau}\}(1 + o(1)) \qquad (4.413)$$

Hence

$$\mathrm{Res}_{z=p_1}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) + \mathrm{Res}_{z=p_2}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right)$$

$$= \frac{N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)}{\mathrm{Re}\{a_{1,\tau}\}}\frac{1}{|P_1|^2|P_1 - P_1^*|^2 \displaystyle\prod_{k,p_k\in\mathscr{P}_{k,\mathbb{R}}}|P_1 - P_k|^2}$$

$$\times \frac{(1 + o(1))}{\prod_{k,p_k\in\mathscr{P}_{k,\mathbb{C}}}|P_1 - P_k|^2|P_1^* - P_k|^2}. \qquad (4.414)$$

The convergence rate of the contribution is determined by the factor

$$\frac{N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)}{2\mathrm{Re}\{a_{1,\tau}\}}. \qquad (4.415)$$

4.5.1.6   Upper Estimation of $\overline{C}(\theta,\tau)$

Asymptotic behaviors of the numerator and the denominator of $\overline{C}(\theta,\tau)$ are presented in the previous subsections. Using the results, this subsection is devoted to treat an upper estimation of $\overline{C}(\theta,\tau)$. Equations (4.400) and (4.414) describe the asymptotic behavior of the denominator of $\overline{C}(\theta,\tau)$. Equation (4.376) gives an upper bound for the absolute value of the numerator of $\overline{C}(\theta,\tau)$.

$$|\overline{C}(\theta,\tau)| \leq \sum_{k,p_k\in\mathscr{P}_\mathbb{T}}\sum_{m,p_m\in\mathscr{P}_\mathbb{T}} C_{a,k}C_{a,m}\prod_{l,Z_l=P_k}(|a_{k,\tau}| + |b_{l,\tau}|)\prod_{l,Z_l=P_m}(|a_{m,\tau}| + |b_{l,\tau}|)$$

$$\times \frac{1}{\displaystyle\sum_{k,p_k\in\mathcal{P}_{\mathbb{T}}} \frac{N(p_k,\theta)N\left(\dfrac{1}{p_k},\theta\right)}{\mathrm{Re}\{a_{k,\tau}\}}+1}\frac{K_{N,k,m}}{K_D}(1+o(1)) \tag{4.416}$$

where

$$K_{N,k,m}=4\frac{\displaystyle\prod_{l,Z_l\neq P_k}|P_k-Z_l|\displaystyle\prod_{l,Z_l\neq P_m}|P_m-Z_l|}{\displaystyle\prod_{l=1,l\neq k}^{n_\alpha}|P_k-P_l|\displaystyle\prod_{l=1,l\neq m}^{n_\alpha}|P_m-P_l|} \tag{4.417}$$

and if P_1 is real,

$$K_D=\frac{1}{2\displaystyle\prod_{k,p_k\in\mathcal{P}_{k,\mathbb{R}}}|P_1-P_k|^2\displaystyle\prod_{k,p_k\in\mathcal{P}_{k,\mathbb{C}}}|P_1-P_k|^2|P_1-P_k^*|^2} \tag{4.418}$$

if P_1 is not real,

$$K_D=\frac{1}{2\left(\mathrm{Im}\{P_1\}\right)^2\displaystyle\prod_{k,p_k\in\mathcal{P}_{k,\mathbb{R}}}|P_1-P_k|^2\displaystyle\prod_{k,p_k\in\mathcal{P}_{k,\mathbb{C}}}|P_1-P_k|^2|P_1^*-P_k|^2}. \tag{4.419}$$

In order to construct an upper estimation of $\overline{C}(\theta,\tau)$, the following cases have to be discussed.

- If $\|\widehat{H}(z,\theta)\|_2^2\to\infty$ as $\tau\to\infty$, then the residue at p_k calculated in (4.400) or (4.414) converges to infinity. We assume that

$$k>1\ \Rightarrow\ \frac{a_{k,\tau}}{\prod_{z_l\in\mathcal{Z}_{s,k}}|b_{l,\tau}|^2}=O\left(\frac{a_{1,\tau}}{\prod_{z_l\in\mathcal{Z}_{s,1}}|b_{l,\tau}|^2}\right), \tag{4.420}$$

i.e. p_1 is has one of the fastest contribution in $\|\widehat{H}(z,\theta)\|_2^2$ from the set $\mathcal{P}_{\mathbb{T}}$. Therefore, assumption $\|\widehat{H}(z,\theta)\|_2^2\to\infty$ yields $\frac{\prod_{z_l\in\mathcal{Z}_{s,1}}|b_{l,\tau}|^2}{a_{1,\tau}}\to\infty$. As we concluded during the investigation of the factor $N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)$ in §4.5.1.3, $\mathcal{Z}_{e,1}$ and $\mathcal{Z}_{f,1}$ are empty sets which are defined at the beginning of §4.5.1.3. Moreover, we known that

$$a_{1,\tau}=o\left(\prod_{z_l\in\mathcal{Z}_{s,1}}|b_{l,\tau}|^2\right). \tag{4.421}$$

Therefore

$$N(p_1,\theta)N\left(\frac{1}{p_1},\theta\right)=\prod_{z_l\in\mathcal{Z}_{s,1}}|b_{l,\tau}|^2\prod_{z_l\in\mathcal{Z}_{o,1}}|P_1-Z_l|^2(1+o(1)) \tag{4.422}$$

and substituting in the corresponding factor of (4.376)

$$\prod_{l,Z_l=P_1}(|a_{1,\tau}|+|b_{l,\tau}|) = \prod_{l,Z_l=P_1}|b_{l,\tau}|(1+o(1)). \tag{4.423}$$

Therefore

$$\overline{C}(\theta,\tau) = O\left(\frac{\left(\prod_{l,Z_l=P_1}|b_{l,\tau}|\right)^2}{\prod_{z_l\in\mathcal{Z}_{s,1}}|b_{l,\tau}|^2\frac{1}{\mathrm{Re}\{a_{1,\tau}\}}}\right) = O\left(\mathrm{Re}\{a_{1,\tau}\}\right). \tag{4.424}$$

- $\| \widehat{H}(z,\theta) \|_2^2$ is bounded as $\tau \to \infty$. It means that only the numerator of $\overline{C}(\theta,\tau)$ determines the convergence rate. An equivalent assumption is that for all $p_k \in \mathcal{P}_{\mathbb{T}}$ the corresponding factors of $N(p_k,\theta)N\left(\frac{1}{p_k},\theta\right)$ are bounded

$$\forall p_k \in \mathcal{P}_{\mathbb{T}} : \frac{\displaystyle\prod_{l,Z_l=P_k}(b_{l,\tau}-a_{k,\tau})(b_{l,\tau}+a_{k,\tau})}{a_{k,\tau}} = O(1). \tag{4.425}$$

If $\mathcal{Z}_{e,k}$ and $\mathcal{Z}_{f,k}$ are empty sets then using (4.386) we have

$$\prod_{l,Z_l=P_k}(b_{l,\tau}-a_{k,\tau})(b_{l,\tau}+a_{k,\tau}) = \prod_{l,Z_l=P_k}b_{l,\tau}^2(1+o(1)). \tag{4.426}$$

From note (4.2.10) we conclude that

$$\prod_{l,Z_l=P_k}b_{l,\tau} = O(\sqrt{|a_{k,\tau}|}) \text{ or } N(p_1) = O(\sqrt{|a_{k,\tau}|}). \tag{4.427}$$

If $z_l \in \mathcal{Z}_{e,1}\cup\mathcal{Z}_{f,1}$, then $|a_{1,\tau}|+|b_{l,\tau}| = O(a_{1,\tau})$. Therefore

$$\prod_{l,Z_l=P_k}(|a_{k,\tau}|+|b_{l,\tau}|) = O(|a_{1,\tau}|) = O(\sqrt{|a_{k,\tau}|}) \tag{4.428}$$

where in the last equation we used the fact that $a_{k,\tau} \to 0$. Factors from (4.376) can be estimated as

$$\left(\prod_{l,Z_l=P_k}(|a_{k,\tau}|+|b_{l,\tau}|)\right)\left(\prod_{l,Z_l=P_m}(|a_{m,\tau}|+|b_{l,\tau}|)\right) = O(a_{1,\tau}) \tag{4.429}$$

First, let us assume that θ_* determines a pole for which $a_{1,\tau}$ converges to 0 at a rate faster than $\frac{1}{\tau}$, i.e. $a_{1,\tau} = o\left(\frac{1}{\tau}\right)$. In this case in (4.376) the limit of $(1-|a_{1,\tau}|)^{\tau-1}$ is a positive number. In both cases above $\overline{C}(\theta_*,\tau) = O(a_{1,\tau}) = o\left(\frac{1}{\tau}\right)$. This is a contradiction because in §4.4.5 an example was presented in which the convergence rate of $\overline{C}(\theta_*,\tau)$ was $\frac{1}{\tau}$.

Second, $|a_{1,\tau}|\tau \to \infty$ is supposed. The convergence rate of $\overline{C}(\theta_*, \tau)$ is determined by

$$(1 - |a_{1,\tau}|)^{2(\tau-1)} o(|a_{1,\tau}|). \tag{4.430}$$

Note 4.2.8 shows that in this case $\overline{C}(\theta_*, \tau)$ converges to zero at exponential rate. This is a contradiction because of the same reason as in the previous case.

In both cases the overall convergence rate of $\overline{C}(\theta_*, \tau)$ is less than $\frac{1}{\tau}$. An example construction of §4.4.5 results $\frac{1}{\tau}$ as convergence rate of $\overline{C}(\theta, \tau)$. This is a contradiction because θ_* was supposed to be an optimal parameter vector for all τ.

The ultimate conclusion of this section is that with assumptions 4.5.1 and 4.5.2 the convergence rate of $\overline{C}(\theta_*, \tau)$ is $\frac{1}{\tau}$.

4.5.2 Handling Unstable Poles

Until this point in the generalized case it has been proved that for an optimal parameter vector sequence θ_*

$$\overline{C}(\theta_*, \tau) \sim \frac{K}{\tau} \tag{4.431}$$

where K is an appropriate, non-zero constant. Moreover, it has been also proved that every pole p_k converging to $\mathbb{T}$ can be represented as

$$p_k = P_k(1 - a_{k,\tau}) \tag{4.432}$$

where $P_k \in \mathbb{T}$ and $a_{k,\tau} = O\left(\frac{1}{\tau}\right)$. It means that be increasing the convergence rate of $a_{k,\tau}$ the cost function cannot be decreased. Now we prove that using assumptions 4.5.1 and 4.5.2, the cost function can be decreased by replacing unstable poles by stable ones.

In this part of the thesis we study the effect of inverting unstable poles which converge to $\mathbb{T}$. The numerator and the denominator of $\overline{C}(\theta, \tau)$ are separately examined. It will be proved that inverting a pole converging to $\mathbb{T}$ does not change the asymptotic value of $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$ but may increase the asymptotic value of the numerator. Since simply inverting a pole may not cause an increase of the asymptotic value of $\left|\int_I T^*(\Omega)e^{i\omega\tau}\widehat{H}(\Omega, \theta)d\omega\right|$, an additional step is required. A phase correction can be applied which does not change the absolute value of the asymptotic contribution in the numerator of $\overline{C}(\theta, \tau)$, but does change signs of the corresponding terms in $T_2(\theta, \tau)$.

The phase correction sequence is denoted by $c_{k,\tau}$ where k denotes the index of the pole this correction sequence belongs to.

$$\tilde{p}_k = p_k e^{ic_{k,\tau}} = P_k(1 - a_{k,\tau}e^{i\tilde{c}_{k,\tau}}) \tag{4.433}$$

where $c_{k,\tau} \to 0$, $\tilde{c}_{k,\tau} \to 0$, $c_{k,\tau} = O(a_{k,\tau})$, $\tilde{c}_{k,\tau} = O(a_{k,\tau})$ and $\forall \tau : c_{k,\tau} \in \mathbb{R}$. In order to preserve the asymptotic value of $N(p_k, \theta)$, i.e. the numerator of $\text{Res}_{z=p_k}\widehat{H}(z, \theta)$ a subset of zeros is transformed, too. The zeros, which converge to P_k at least at rate of $a_{k,\tau}$ is shifted such that their contribution remains the same. In the following steps of the proof a sequence of transformations is applied for unstable p_k in order to construct a new configuration of

poles and zeros. This configuration determines a new θ for which $\overline{C}(\theta, \tau) > \overline{C}(\theta_*, \tau)$, if τ is sufficiently large.

In this section both assumptions 4.5.1 and 4.5.2 are applied.

4.5.2.1   Behavior on $\int_I |\widehat{H}(\Omega, \theta)|^2 d\omega$

Effects of the pole inversion are studied in every term of the following sum:

$$\frac{1}{2\pi} \int_I |\widehat{H}(\Omega, \theta)|^2 d\omega = \sum_{k, p_k \in \mathcal{P}} \text{Res}_{z=p_k} \widehat{H}(z, \theta) \frac{1}{z} \widehat{H}\left(\frac{1}{z}, \theta\right). \tag{4.434}$$

Since we had assumed that every pole of the optimal parameter vector θ_* converges to a unique point of $\mathbb{T}$, for all τ sufficiently large, we have

$$\sum_{k, p_k \in \mathcal{P}} \text{Res}_{z=p_k} \frac{N(z, \theta) N\left(\frac{1}{z}, \theta\right)}{z(z - p_k)\left(\frac{1}{z} - p_k\right) \prod\limits_{\substack{k=1 \\ k \neq l}}^{n_\alpha} (z - p_l)\left(\frac{1}{z} - p_l\right)}$$

$$= \sum_{k, p_k \in \mathcal{P}} \frac{N(p_k, \theta) N\left(\frac{1}{p_k}, \theta\right)}{p_k\left(\frac{1}{p_k} - p_k\right) \prod\limits_{\substack{k=1 \\ k \neq l}}^{n_\alpha} (p_k - p_l)\left(\frac{1}{p_k} - p_l\right)}. \tag{4.435}$$

It is worth noting that in the last step assumption 4.5.1 is used because no multiple pole appears in (4.435).

Consider p_k with $k \neq l$. Then the factor in the denominator

$$(p_k - p_l)\left(\frac{1}{p_k} - p_l\right) \tag{4.436}$$

is invariant when replacing p_k by $\frac{1}{p_k}$. And using again assumption 4.5.1 limit of these factors is non-zero. Similarly, $N(p_k, \theta) N\left(\frac{1}{p_k}, \theta\right)$ remains the same when replacing p_k by $\frac{1}{p_k}$. A pole p_k can be written as

$$p_k = P_k(1 + a_{k, \tau}) \text{ with } \begin{cases} P_k = \pm 1, & \text{in the real case,} \\ P_k = e^{j\phi_k}, & \text{in the complex case.} \end{cases} \tag{4.437}$$

If p_k is real, then

$$z(z - p_k)\left(\frac{1}{z} - p_k\right) = (p_k - z)\left(z - \frac{1}{p_k}\right) p_k. \tag{4.438}$$

Since $|p_k| > 1$, $\frac{1}{p_k}$ is inside the unit circle. Substituting $z = \frac{1}{p_k}$,

$$\left(p_k - \frac{1}{p_k}\right)p_k = (p_k^2 - 1) = ((1 + a_{k,\tau})^2 - 1) = 2a_{k,\tau}(1 + O(a_{k,\tau})). \tag{4.439}$$

Inverting p_k means that

$$\tilde{p}_k = \frac{1}{p_k} = P_k(1 - a_{k,\tau}) + O(a_{k,\tau}). \tag{4.440}$$

The new contribution to the denominator of (4.435) is

$$z(z - \tilde{p}_k)\left(\frac{1}{z} - \tilde{p}_k\right). \tag{4.441}$$

Substituting $z = \tilde{p}_k$ leads to

$$\tilde{p}_k\left(\frac{1}{\tilde{p}_k} - \tilde{p}_k\right) = (1 - \tilde{p}_k^2) = 2a_{k,\tau}(1 + O(a_{k,\tau})). \tag{4.442}$$

Therefore

$$\text{Res}_{z=p_k}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) \sim \text{Res}_{z=\tilde{p}_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right) \tag{4.443}$$

where $\tilde{\theta}$ is constructed from θ such that p_k is replaced by $\tilde{p}_k$.

Assumption 4.5.1 says that for all k sufficiently large, the multiplicity of a pole is one. Therefore, in the case of a real pole multiplying by an arbitrary phase correction sequence defined in (4.433) may result in a complex pole. Hence, in the case of a real pole the phase changing is not taken into consideration.

If p_k is complex then

$$z(z - p_k)(z - p_k^*)\left(\frac{1}{z} - p_k\right)\left(\frac{1}{z} - p_k^*\right)$$

$$= \frac{1}{z}(z - p_k)(z - p_k^*)|p_k|^2\left(z - \frac{1}{p_k}\right)\left(z - \frac{1}{p_k^*}\right). \tag{4.444}$$

The reciprocals $\frac{1}{p_k}$ and $\frac{1}{p_k^*}$ are inside the unit circle, so the residues have to be calculated at $\frac{1}{p_k}$ and $\frac{1}{p_k^*}$. Substituting $z = \frac{1}{p_k}$

$$p_k\left(\frac{1}{p_k} - p_k\right)\left(\frac{1}{p_k} - p_k^*\right)|p_k|^2\left(\frac{1}{p_k} - \frac{1}{p_k^*}\right)$$

$$= \left(\frac{1}{p_k} - p_k\right)(1 - |p_k|^2)(p_k^* - p_k).$$

Applying $p_k = P_k(1 + a_{k,\tau})$ and $p_k^* = P_k^*(1 + a_{k,\tau}^*)$, this is equal to

$$(P_k^* - P_k + o(1))(-2\mathrm{Re}\{a_{k,\tau}\} + O(|a_{k,\tau}|^2))(P_k^* - P_k + o(1))$$

$$= |P_k - P_k^*|^2 2\mathrm{Re}\{a_{k,\tau}\}(1 + o(1)). \tag{4.445}$$

Inverting p_k and p_k^* means that

$$\tilde{p}_k = \frac{1}{p_k} = P_k^*(1 - a_{k,\tau}^*) + O(|a_{k,\tau}|^2), \tag{4.446}$$

and

$$\tilde{p}_k^* = \frac{1}{p_k^*} = P_k(1 - a_{k,\tau}) + O(|a_{k,\tau}|^2). \tag{4.447}$$

Substituting $z = \tilde{p}_k$ into

$$z(z - \tilde{p}_k^*)\left(\frac{1}{z} - \tilde{p}_k\right)\left(\frac{1}{z} - \tilde{p}_k^*\right)$$

we get

$$\tilde{p}_k(\tilde{p}_k - \tilde{p}_k^*)\left(\frac{1}{\tilde{p}_k} - \tilde{p}_k\right)\left(\frac{1}{\tilde{p}_k} - \tilde{p}_k^*\right) = (\tilde{p}_k - \tilde{p}_k^*)\left(1 - |\tilde{p}_k|^2\right)\left(\frac{1}{\tilde{p}_k} - \tilde{p}_k\right)$$

$$= (P_k^* - P_k + o(1))(2\mathrm{Re}\{a_{k,\tau}\} + O(|a_{k,\tau}|^2))(P_k - P_k^* + o(1))$$

$$= |P_k - P_k^*|^2 2\mathrm{Re}\{a_{k,\tau}\}(1 + o(1)). \tag{4.448}$$

Therefore

$$\mathrm{Res}_{z=p_k}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) \sim \mathrm{Res}_{z=\tilde{p}_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right),$$

$$\mathrm{Res}_{z=p_k^*}\widehat{H}(z,\theta)\frac{1}{z}\widehat{H}\left(\frac{1}{z},\theta\right) \sim \mathrm{Res}_{z=\tilde{p}_k^*}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right)$$

where $\tilde{\theta}$ is constructed from θ such that p_k is replaced by $\tilde{p}_k$ and p_k^* is replaced by $\tilde{p}_k^*$.

4.5.2.2 Phase Correction

An arbitrary phase change can be applied if complex conjugate pairs are treated. The conjugate pairs are modified together, i.e.

$$\tilde{p}_k = p_k e^{ic_{k,\tau}}, \tilde{p}_k^* = p_k^* e^{-ic_{k,\tau}}. \tag{4.449}$$

The presented phase correction can be changed $N(p_k, \theta)$ significantly. For example, let us assume that we have a pole p_1 with $P_1 = i$ and $a_{1,\tau} = \frac{1}{\tau}$ and a zero converging to i with $b_{1,\tau} = \frac{1}{\tau} - \frac{1}{\tau^2}e^{i/\tau^2}$. If the phase correction is chosen as $\tilde{c}_{k,\tau} = \frac{1}{\tau^2}e^{i/\tau^2}$, then $\forall\tau : N(\tilde{p}_k, \theta) = 0$, but $\forall\tau N(p_k, \theta) \neq 0$. In order to overcome this problem not only the phase of p_k but zeros from the set $\mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$ are transformed, too.

Our aim is to modify the elements of the set $\mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$ such that

$$\lim_{\tau\to\infty} \frac{N(p_k, \theta)}{N(\tilde{p}_k, \theta)} = 1. \tag{4.450}$$

Let us denote $q = p_k - \tilde{p}_k$. $\forall z_l \in \mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$ $\tilde{z}_l$ is defined as

$$\tilde{z}_l = z_l + q. \tag{4.451}$$

$\tilde{\theta}$ is constructed from θ such that p_k is replaced by $\tilde{p}_k$, $\forall z_l \in \mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$ are replaced by $\tilde{z}_l$ and every other pole and zero remain unchanged. First the denominator of $\widehat{H}(z, \tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z}, \tilde{\theta}\right)$ is examined, then we focus on its numerator.

We assume that $\forall\tau : \text{Im}\{p_1(\tau)\} \neq 0$ and that $\forall\tau : p_2(\tau) = p_1^*(\tau)$. In the residue at $p_l \in \mathcal{P}_{\mathbb{T}}$, $l > 2$ the contribution of p_1 can be written as

$$|p_l - \tilde{p}_1|^2|p_l^* - \tilde{p}_1|^2 = |p_l - p_1 e^{ic_{1,\tau}}|^2|p_l^* - p_1 e^{ic_{1,\tau}}|^2 \tag{4.452}$$

$$= |p_l - p_1(1 + ic_{1,\tau} + O\left(|c_{1,\tau}|^2\right))|^2|p_l^* - p_1(1 + ic_{1,\tau} + O\left(|c_{1,\tau}|^2\right))|^2 \tag{4.453}$$

where in the last step the usual Taylor expansion of the exponential function is applied. In can be done without assumptions because the exponential function is regular on the whole complex plane. Since $c_{k,\tau} = o(1)$ we have

$$|p_l - \tilde{p}_1|^2|p_l^* - \tilde{p}_1|^2 = |P_l - P_1|^2|P_l^* - P_1|^2(1 + o(1)) \tag{4.454}$$

which is asymptotically the same as the contribution of $|p_l - p_1|^2|p_l^* - p_1|^2$. In the denominator of $\widehat{H}(z, \tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z}, \tilde{\theta}\right)$ we have

$$z(z - \tilde{p}_1)\left(\frac{1}{z} - \tilde{p}_1\right)(z - \tilde{p}_1^*)\left(\frac{1}{z} - \tilde{p}_1^*\right). \tag{4.455}$$

In order to calculate the residue at $z = \tilde{p}_1$, the following expression is evaluated:

$$z\left(\frac{1}{z} - \tilde{p}_1\right)(z - \tilde{p}_1^*)\left(\frac{1}{z} - \tilde{p}_1^*\right)\Bigg|_{z=\tilde{p}_1} \tag{4.456}$$

Substituting $z = \tilde{p}_1$ leads to

$$= \tilde{p}_1\left(\frac{1}{\tilde{p}_1} - \tilde{p}_1\right)(\tilde{p}_1 - \tilde{p}_1^*)\left(\frac{1}{\tilde{p}_1} - \tilde{p}_1^*\right)$$

$$= p_1 e^{ic_{1,\tau}} \left(\frac{1}{p_1 e^{ic_{1,\tau}}} - p_1 e^{ic_{1,\tau}} \right) (p_1 e^{ic_{1,\tau}} - p_1^* e^{-ic_{1,\tau}}) \left(\frac{1}{p_1 e^{ic_{1,\tau}}} - p_1^* e^{-ic_{1,\tau}} \right)$$

$$= (P_1^* - P_1 + O(|a_{1,\tau}| + c_{1,\tau})) (P_1 - P_1^* + O(|a_{1,\tau}| + c_{1,\tau})) (1 - |p_1|^2)$$

$$= 2\mathrm{Re}\{a_{1,\tau}\}|P_1 - P_1^*|^2(1 + o(1)). \tag{4.457}$$

Therefore one can conclude that inverting an unstable pole that converges to $\mathbb{T}$ or modifying the phase of a complex pole by multiplying it with a complex sequence converging to 1 does not change the asymptotic value of the denominator of $\mathrm{Res}_{z=p_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right)$.

In the following, the numerator of $\mathrm{Res}_{z=p_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right)$ is investigated. Two cases are treated: whether the norm of $\| H(z,\theta) \|_2^2$ converges to infinity, or not.

- $\| \widehat{H}(z,\theta) \|_2^2 \to \infty$. In this case, $\widehat{H}(z,\theta)$ contains only zeros from the set $\mathcal{Z}_{o,k} \cup \mathcal{Z}_{s,k}$. Since for all $z_l \in \mathcal{Z}_{o,k}$

$$(p_k - z_l) = |P_k - Z_l|(1 + o(1)) \tag{4.458}$$

and

$$\left(\frac{1}{p_k} - z_l \right) = |P_k^* - Z_l|(1 + o(1)) \tag{4.459}$$

the limit value is independent of the phase transformation. If $z_l \in \mathcal{Z}_{s,k}$, then

$$p_k - z_l = P_k(-a_{k,\tau} + b_{l,\tau}) = P_k b_{l,\tau}(1 + o(1)) \tag{4.460}$$

or

$$\frac{1}{p_k} - z_l^* = P_k^*(-a_{k,\tau}^* + b_{l,\tau}^*) = P_k b_{l,\tau}^*(1 + o(1)). \tag{4.461}$$

The asymptotic contribution of these factors does not depend on the defined phase transformation. From §4.4.5 and §4.5.1 we known that both p_k and $\tilde{p}_k$ converge to $\mathbb{T}$ at the rate $\frac{1}{\tau}$. Hence the asymptotic contribution of $\tilde{p}_k$ is also independent of $a_{k,\tau}$. Therefore, we can conclude that if $\| \widehat{H}(z,\theta) \|_2^2 \to \infty$ then the presented phase transformation does not change the asymptotic value of the numerator of $\mathrm{Res}_{z=p_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right)$.

- $\| \widehat{H}(z,\theta) \|_2^2$ remains bounded. In this case we know that $\mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k} \neq \emptyset$, it is possible that the set $\mathcal{Z}_{o,k} \cup \mathcal{Z}_{s,k}$ is not empty. Then the same argument like in the previous case can be applied. Using assumption 4.5.1, it can be shown that

$$\mathrm{Res}_{z=p_k}\widehat{H}(z,\tilde{\theta})\frac{1}{z}\widehat{H}\left(\frac{1}{z},\tilde{\theta}\right) \to 0. \tag{4.462}$$

In the corresponding factors of the denominator we have

$$z(z - p_k^*) \left(\frac{1}{z} - p_k \right) \left(\frac{1}{z} - p_k^* \right),$$

(4.463)

substituting $z = p_k$ we get

$$p_k(p_k - p_k^*) \left(\frac{1}{p_k} - p_k \right) \left(\frac{1}{p_k} - p_k^* \right).$$

(4.464)

Only the factor $\left(\frac{1}{p_k} - p_k^* \right) \to 0$ and its convergence rate is

$$\left(\frac{1}{p_k} - p_k^* \right) = O(a_{k,\tau}).$$

(4.465)

In the numerator $z_l \in \mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$

$$(p_k - z_l) \left(\frac{1}{p_k} - z_l \right) (p_k - z_l^*) \left(\frac{1}{p_k} - z_l^* \right)$$

$$= |P_k|^2 (1 + o(1))(-a_{k,\tau} - b_{k,\tau})(a_{k,\tau}^* + b_{k,\tau}^*)$$

(4.466)

and by the definition of $\mathcal{Z}_{e,k}$ and $\mathcal{Z}_{f,k}$ $b_{k,\tau} = O(a_{k,\tau})$ we conclude that

$$N(p_k, \theta) N \left(\frac{1}{p_k}, \theta \right) = O(|a_{k,\tau}|^2).$$

(4.467)

(4.467) and (4.465) proves (4.462).

Assumption 4.5.2 and the result form §4.4.5 guarantee that the same is true for residues at all the poles of $\widehat{H}(z, \tilde{\theta}) \frac{1}{z} \widehat{H} \left(\frac{1}{z}, \tilde{\theta} \right)$, except residue at 0. Therefore,

$$\| \widehat{H}(z, \theta) \|_2^2 \to 1$$

(4.468)

as $\tau \to \infty$.

4.5.2.3   Effect on $\left| \int_I T^*(\Omega) e^{i\omega\tau} \widehat{H}(\Omega, \theta) d\omega \right|$

Inverting an unstable pole that converges to $\mathbb{T}$ may have significant effect on the asymptotic value of $\left| \int_I T^*(\Omega) e^{i\omega\tau} \widehat{H}(\Omega, \theta) d\omega \right|$. Similarly, the phase correction has asymptotic importance. Like in the previous subsection, first the effect of the inversion of an unstable pole is investigated, then we study the phase correction transformation.

In p_k is unstable than the function

$$T(z^{-1}) z^{\tau-1} \widehat{H}(z, \theta)$$

(4.469)

has no pole at $\frac{1}{p_k}$ inside the unit circle. It means that inverting an unstable pole increases the number of the terms in the residual form of $\int_I T^*(\Omega)e^{i\omega\tau}\widehat{H}(\Omega,\theta)d\omega$. It was seen in §4.5.1.2, in the numerator of $\overline{C}(\theta,\tau)$ the contribution of the residue at p_k is determined by the factor $N(p_k)$ and the factor $(1-|a_{k,\tau}|)^\tau$. Depending on the limit of $\int_I |\widehat{H}(\Omega,\theta)|^2 d\omega$ the convergence rate of $N(p_k,\theta)$ is determined by the factor

$$\prod_{l,Z_l=P_k} (b_{l,\tau} - a_{k,\tau}). \tag{4.470}$$

In the special case when $a_{k,\tau} = o(\prod_{l,Z_l=P_k} b_{l,\tau})$, i.e. $\mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k} = \emptyset$ we have

$$\prod_{l,Z_l=P_k} (b_{l,\tau} - a_{k,\tau}) = \prod_{l,Z_l=P_k} b_{l,\tau}(1 + o(1)). \tag{4.471}$$

Let us define $C_{a,k}$ as

$$C_{a,k} = (1 - a_{k,\tau})^\tau \tag{4.472}$$

which generalists the definition of (4.220).

Now the sign of the contribution of a stable pole p_k that converges to $\mathbb{T}$ will be studied.

In most cases modifying the phase of a real pole causes that it becomes a complex pole. Polynomials which occur in this thesis have real coefficients, thus modifying only the phase of just one real pole leads to complex coefficients. Therefore, for real poles the phase modification is not allowed.

In the case of complex poles, the phase transformation of p_k and p_k^* (4.449) are applied simultaneously. Assuming that $p_2 = p_1^*$, the contribution of the complex conjugate pole pair is

$$\mathrm{Res}_{z=p_1} T(z^{-1})z^{\tau-1}H(z,\theta) + \mathrm{Res}_{z=p_2} T(z^{-1})z^{\tau-1}\widehat{H}(z,\theta)$$

$$= 2\mathrm{Re}\left\{\mathrm{Res}_{z=p_1} T(z^{-1})z^{\tau-1}H(z,\theta)\right\}$$

$$= 2\mathrm{Re}\left\{ T\left(\frac{1}{P_1}\right) p_1^{\tau-1} \frac{\prod_{l,Z_l=P_1}(p_1 - z_l)\prod_{l,Z_l\neq P_1}(p_1 - z_l)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)}(1 + o(1))\right\}. \tag{4.473}$$

Assuming that $p_1 = P_1(1 - |a_{1,\tau}|)e^{i\phi_{1,\tau}}$, we have

$$2\mathrm{Re}\left\{\mathrm{Res}_{z=p_1} T(z^{-1})z^{\tau-1}H(z,\theta)\right\} =$$

$$= (1 - |a_{1,\tau}|)^{\tau-1} 2\mathrm{Re}\left\{ T\left(\frac{1}{P_1}\right) P_1 e^{i\phi_{1,\tau}(\tau-1)} \frac{N_s(p_1)N_f(p_1)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)}(1 + o(1))\right\}. \tag{4.474}$$

where

$$N_s(p_1) = \prod_{l,z_l \in \mathcal{Z}_{o,1} \cup \mathcal{Z}_{s,1}} (p_1 - z_l) \tag{4.475}$$

and

$$N_f(p_1) = \prod_{l,z_l \in \mathcal{Z}_{e,1} \cup \mathcal{Z}_{f,1}} (p_1 - z_l). \tag{4.476}$$

$$2\mathrm{Re}\left\{\mathrm{Res}_{z=p_1} T(z^{-1}) z^{\tau-1} H(z,\theta)\right\} =$$

$$= 2C_{a,k}(1 + o(1))\left(\mathrm{Re}\left\{e^{i\phi_{1,\tau}(\tau-1)}\right\}\mathrm{Re}\left\{T\left(\frac{1}{P_1}\right) P_1 \frac{N_s(p_1)N_f(p_1)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)}\right\}\right.$$

$$\left. -\mathrm{Im}\left\{e^{i\phi_{1,\tau}(\tau-1)}\right\}\mathrm{Im}\left\{T\left(\frac{1}{P_1}\right) P_1 \frac{N_s(p_1)N_f(p_1)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)}\right\}\right) \tag{4.477}$$

Notes 4.2.11 and 4.2.12 ensure that there exists a $\tilde{p}_1$ such that

- $|p_1| = |\tilde{p}_1|$,

- $\lim_{\tau\to\infty} p_1 = \lim_{\tau\to\infty} \tilde{p}_1$,

- $|p_1 - \tilde{p}_1| < \frac{1}{\tau}$

and the signs of $\mathrm{Re}\left\{e^{i\phi_{1,\tau}(\tau-1)}\right\}$ and $\mathrm{Im}\left\{e^{i\phi_{1,\tau}(\tau-1)}\right\}$ are arbitrary. Therefore $\tilde{p}_1$ is nothing else than result of a phase transformation step. p_2 is modified such that $p_2^* = p_1$. The limits remain the same, so the factor $(P_1 - P_2)$ in the denominator of (4.477) is invariant with respecting to the introduced phase transformation step. The corresponding zeros of $N_f(z)$ are modified according to the presented transformation in (4.451). It is clear that

$$N_s(p_1) = \prod_{l,z_l \in \mathcal{Z}_{o,1} \cup \mathcal{Z}_{s,1}} (p_1 - z_l) = \prod_{l,z_l \in \mathcal{Z}_{o,1} \cup \mathcal{Z}_{s,1}} (\tilde{p}_1 - \tilde{z}_l) = \tilde{N}_s(\tilde{p}_1) \tag{4.478}$$

By the definition of $\tilde{p}_1$ we have

$$T\left(\frac{1}{P_1}\right) P_1 \frac{N_s(p_1)N_f(p_1)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)} = T\left(\frac{1}{P_1}\right) P_1 \frac{N_s(\tilde{p}_1)N_f(\tilde{p}_1)}{\prod_{l=2}^{n_\alpha}(P_1 - P_l)}. \tag{4.479}$$

hence for all stable p_k

$$\left|\mathrm{Res}_{z=p_k} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\theta)\right| \sim \left|\mathrm{Res}_{z=\tilde{p}_k} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\tilde{\theta})\right| \tag{4.480}$$

or

$$\left|\mathrm{Res}_{z=p_k} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\theta) + \mathrm{Res}_{z=p_k^*} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\tilde{\theta})\right| \tag{4.481}$$

$$\sim \left| \mathrm{Res}_{z=\tilde{p}_k} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\theta) + \mathrm{Res}_{z=\tilde{p}_k^*} T\left(\frac{1}{z}\right) z^{\tau-1} \widehat{H}(z,\tilde{\theta}) \right|. \tag{4.482}$$

Hence we conclude that for all p_k which converge to $\mathbb{T}$ and for all $z_l \in \mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$, there exists a $\tilde{p}_k$ and $\tilde{z}_l$ such that the absolute value of their asymptotic contribution remains the same but the signs of their contributions are arbitrary. We refer to this transformation of the poles and zeros as the phase correction step.

4.5.2.4   Evaluating $\overline{C}(\theta,\tau)$

The asymptotic properties of the numerator and denominator of $\overline{C}(\theta,\tau)$ are discussed in §4.5.2.3 and §4.5.2.1. Two kinds of transformation step are treated:

- inverting an unstable pole converging to $\mathbb{T}$,

- multiplying a complex conjugate pole pair by a phase correction sequence.

Using the composition of the sequence $\theta(\tau)$ introduced in §4.4.3, indirect technique is applied to prove that asymptotically an optimal parameter vector sequence $\theta_*(\tau)$ does not contain any unstable pole. This means that, it is proved that for all τ sufficiently large for every $\theta(\tau)$ which contains an unstable pole subsequence a $\tilde{\theta}(\tau)$ can be constructed such that $\tilde{\theta}(\tau)$ defines the same poles and zeros like $\theta(\tau)$ except the corresponding unstable pole that is inverted. Hence, assuming that a particular optimal parameter vector sequence $\theta_*(\tau)$ containing an unstable pole subsequence, construction of a new $\tilde{\theta}(\tau)$, for which $C(\tilde{\theta}(\tau),\tau) < C(\theta_*(\tau),\tau)$ and $\tilde{\theta}(\tau)$ defines the stabilized pole subsequence, leads to a contradiction. This contradiction proves that asymptotically $\theta_*(\tau)$ does not define any unstable pole subsequence. In the proof assumptions 4.5.1 and 4.5.2 are strongly used.

For a particular θ_* let us assume that $\theta_*(\tau)$ contains unstable pole and let $p_1(\tau),\dots,p_U(\tau)$ denote these unstable poles converging to $\mathbb{T}$. U denotes the number of unstable pole in the corresponding sequence, the set of the pole sequences is denoted by $\mathcal{P}_u = \{p_1(\tau),,\dots,p_U(\tau)\}$. The elements of complex conjugate pole pairs are investigated together. The following algorithm is applied with a starting value $\theta = \theta_*$.

1. Determine $\mathcal{P}_u$ from θ.

2. If $\mathcal{P}_u$ is not empty, then select a real $p_k(\tau)$, or in the case of a complex conjugate pole pair $p_k^*(\tau)$ is also selected. If empty, then the algorithm ends.

3. Invert $p_k(\tau)$, and $p_k^*(\tau)$ in the complex case. The zeros from the set $\mathcal{Z}_{e,k} \cup \mathcal{Z}_{f,k}$ are modified according to (4.451). It guarantees that

$$N(p_k(\tau), \theta_*(\tau)) = N\left(\frac{1}{p_k(\tau)}, \tilde{\theta}(\tau)\right), \tag{4.483}$$

or

$$N(p_k(\tau), \theta_*(\tau)) = N\left(\frac{1}{p_k^*(\tau)}, \tilde{\theta}(\tau)\right), \tag{4.484}$$

$$N(p_k^*(\tau), \theta_*(\tau)) = N\left(\frac{1}{p_k(\tau)}, \tilde{\theta}(\tau)\right). \tag{4.485}$$

Other poles and zeros defined by $\theta_*(\tau)$ are copied into $\tilde{\theta}(\tau)$ without any modification. In §4.5.2.1 it is proved that such transformation does not changed the asymptotic value of the denominator $\overline{C}(\theta_*, \tau)$, i.e.

$$\| \widehat{H}(\Omega, \theta) \|_2^2 \sim \| \widehat{H}(\Omega, \tilde{\theta}) \|_2^2. \tag{4.486}$$

4. Because of assumption 4.5.2 in the numerator of $\overline{C}(\theta, \tau)$ the convergence rate of the residues calculated at p_k and p_k^* is the same like the sum of other residues calculated at the stable poles. However, it is possible that the sign of the contribution of p_k or the sign of the sum of the contributions p_k and p_k^* is different. In this case the phase correction step is applied. If p_k is complex, then new $\tilde{p}_k$ and $\tilde{p}_k^*$ are introduced according to §4.5.2.3. If p_k is real, then for all other stable poles are transformed such that sign of every contribution except the real ones changes.

After this step the number of terms in the numerator of $\overline{C}(\theta, \tau)$ increases. The number of non-zero terms increases and every term has the same sign. Therefore, for all τ sufficiently large,

$$\left| \int_I T^*(\Omega) e^{i\omega\tau} \widehat{H}(\Omega, \theta_*) d\omega \right| < \left| \int_I T^*(\Omega) e^{i\omega\tau} \widehat{H}(\Omega, \tilde{\theta}) d\omega \right|. \tag{4.487}$$

5. p_k or p_k and p_k^* are removed from the set $\mathcal{P}_u$.

6. $\theta = \tilde{\theta}$. Goto to the first step.

It is clear that every time when step 5 is evaluated, for all τ sufficiently large the cost function and the number of unstable poles decrease. Since only finite number of unstable pole is possible, the presented algorithm ends after finite number of steps and results in empty set $\mathcal{P}_u$. The conclusion of §4.5.1 ensures that the cost function cannot be decreased asymptotically by increasing the convergence rate of unstable poles.

The result is that a sequence of parameter vector is constructed $\tilde{\theta}(\tau)$ such that $C(\tilde{\theta}(\tau), \tau) < C(\theta_*(\tau), \tau)$. This a contradiction with the assumption that $\theta_*(\tau)$ is the solution of the optimization problem.

In the presented algorithm does not work in a special case. The situation was already addressed in §4.4.8. In the case when two real pole converge to $\mathbb{T}$ such that their limit points are different, it is possible that inverting of an unstable pole decreases $\overline{C}(\theta, \tau)$. And in this case the phase modifying step cannot be applied. It means that in general the statement which says that for all τ sufficiently large the global minimum of $C(\theta, \tau)$ determines a stable model, is not true. But fortunately, since this is the only exception, the existence theorem is still true. The valid statement says that depending on the target function $T(\Omega)$, for all even τ or for all odd τ sufficiently large the global minimum of $C(\theta, \tau)$ determines a stable system. From the latest statement, the existence theorem can be concluded.

4.6   Weight

In the original cost function (4.1) a weighting function $W(\Omega)$ can be introduced. There are restrictions for the weight function: $W(\Omega)$ must be a meromorphic function on the whole complex plane without poles and zeros on the unit circle (z-domain) nor on the imaginary axes (s-domain). The existence theorem remains true in this case because every step of the proof can be repeated again. Let $W(\omega)$ denote the weighting function. It must be a restriction of a meromorphic function on the whole complex plane $W(z)$. The new cost function equals:

$$C(\theta, \tau) = \| W(z)(T(z)z^{-\tau} - H(z, \theta)) \|_2^2 . \tag{4.488}$$

This is nothing else than a new norm on the space of systems. Introduction this weighting function causes no change in the proof. Every step of the proof can be repeated with a little change that is beside the point.

Chapter 5

Numerical Algorithm

In this chapter a new numerical algorithm is presented for finding automatically a delay where the global minimum of the cost function (4.1) determines a stable model. (In this case the frequency grid I is a finite set.) The idea of the new method is based on the theory of differential equations. Contrary to classical approaches it needs less gradient-like steps during the approximation process. Moreover, we will show examples in which the proposed algorithm returns a better solution than those found in the literature.

First, the problem formulation is presented. Second, an existing method is studied. After that, there is a quick overview of the theory of numerical methods for solving differential equations. Fourth, the new algorithm is introduced and, in the end, numerical and measurement examples are shown.

5.1   Problem Formulation

The original problem is to find the global minimum of the cost function (3.26). Unfortunately, in a computer the continuous frequency grid cannot be handled. Therefore, the DFT defined in (2.7) is used.

The interval I in (3.26) is replaced by a finite set of frequencies. Thus, we introduce the discrete version of the cost function:

$$C(\theta, \tau) = \sum_{k=1}^{F} |T(\Omega_k)e^{-j\omega_k \tau} - H(\Omega_k, \theta)|^2 \tag{5.1}$$

where Ω_k denotes the kth generalized frequency ($\Omega_k = j\omega_k$ in s-domain, $\Omega_k = e^{j\omega_k}$ in z-domain) and F is the number of the frequencies. It should be noted that in (5.1) the meaning of $C(\theta, \tau)$ is redefined. However, it does not cause any problem because throughout this chapter only the new definition is used. Since the Lebesgue integral [Rudin, 1976] contains the discrete measure as a special case, (5.1) is not a new definition, but a special case of (3.26) (of course the factor $2\pi\Delta f$ has to be included). Applying the initiated notations the minimization problem can be written as

$$\theta_* = \arg\min_{\theta} C(\theta, \tau) \text{ where } \tau \text{ is fixed.} \tag{5.2}$$

In order to emphasize that θ_* depends on τ, sometimes the notation $\theta_*(\tau)$ is used.

If the delay value τ is fixed then the minimum of equation (5.1) w.r.t. the parameter vector $\theta_*(\tau)$ has to be found. In the literature there are lots of extensively used and excellent algorithms [Press et al., 2002]. As already mentioned without imposing any constraint it is not guaranteed that the solution $\theta_*(\tau)$ of the minimization problem determines a stable model. The problem of finding the optimal delay can be formulated in the following:

$$\arg\min_\tau C(\theta_*(\tau)) \text{ subject to } H(\Omega, \theta_*(\tau)) \text{ stable.} \tag{5.3}$$

The point is that we have to find a delay value where the minimum of the cost function determines a stable model, and the value of the cost function is minimal subject to this constraint. There is an obvious solution. A new parameter vector is constructed by adding the delay as a new element to θ. Using this new parameter space the minimization problem is

$$[\theta_*, \tau_*] = \arg\min_{[\theta,\tau]} C(\theta, \tau) \text{ subject to } H(\Omega, \theta_*(\tau)) \text{ stable} \tag{5.4}$$

where τ_* denotes the optimal delay value. The problem of this solution is that it imposes a constraint. Without the constraint it can be solved very well with the existing methods [Schoukens and Pintelon, 1991], [Kollár, 1994]. Intuitively, it can be seen that the set of the stable parameter vectors in the whole space is a very sophisticated in spite of the fact that it is a connected subset. But generally the set

$$\{\tau : \theta_*(\tau) \text{ subject to } H(\Omega, \theta_*(\tau)) \text{ is stable}\} \tag{5.5}$$

is not a connected set. An illustration can be seen in figure 5.1. The output of the optimization solver with a fixed delay value τ is denoted by $\theta_{*,\text{os}}(\tau)$. It is important to emphasize that $\theta_{*,\text{os}}(\tau) = \theta_*(\tau)$ cannot be guaranteed. In the figure a Hilbert transformer (see §5.7.1) was the target function, and it was approximated by a 12th order model. Levenberg-Marquardt method (see §5.5.3) was used as minimization procedure. It can be seen in figure 5.1 that the computed extreme values of the complex function are very erratic and several local optima may exist. As we known from the Galois theory of algebra [Jacobson, 1985] it is impossible to determine an expression for the map between a polynomial and its roots if the order is greater than 4. Therefore, the optimization problem (5.3) needs another method.

5.2   The Mean Group Delay Method

One of the possible solutions for problem (5.3) is studied deeply in [Vuerinckx, 1998]. Now we introduce Rudi Vuerinckx's results which was the base of research of this thesis.

[Vuerinckx, 1998] treats the stable IIR filter design problem with additional delay. The main aim of the thesis is to develop an algorithm to find the optimal delay value and the optimal parameter vector in *Chebyshev* sense. This means that the approximation problem has to be solved in the space $L_{\infty,\mathbb{R}}(\mathbb{T})$ (see §2.2). Practical examples showed that
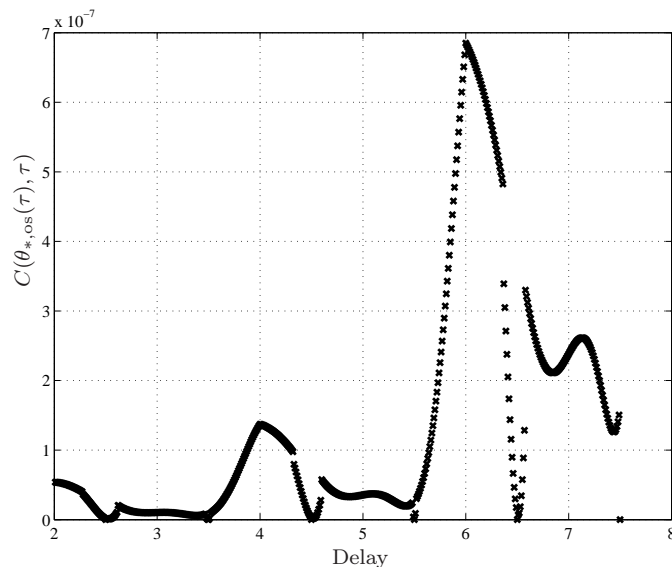
Figure 5.1: Illustration of the complexity of extreme values of the cost function for fixed delay values.

the optimal (stable) delay value in Chebyshev IIR filter approximation is usually close to the optimal (stable) delay value in L_2 IIR filter approximation. Therefore the developed algorithm firstly solves an L_2 approximation problem and then using this as initial value a Chebyshev approximation is performed. For completeness the whole algorithm is introduced, not only the part using the L_2 approximation.

5.2.1 Outline of the Method

An iterative algorithm is used to find the delay for which a stable filter approximation results. In each iteration, an IIR filter is designed at a specific delay, without imposing any stability constraint. If this filter is stable, the solution is found. Otherwise, a suitable delay increment is deduced from the characteristics of this unstable filter and a next iteration is started. The critical part consists of the algorithm used to obtain an appropriate delay increment from the frequency response of the unstable filters: the resulting delay increment must be such that a minimal number of iterations are necessary before reaching a stable filter, while not adding unnecessary delay, because this usually deteriorates the approximation error [Vuerinckx, 1998].

Because it can take many iterations before a stable delay is found, where each iteration involves a full Chebyshev filter design, this method as such can be quite time consuming. To reduce this calculation time, not all filters are designed with the slow Chebyshev method. First, the much faster least squares norm is used to generate the filters. Only after a stable delay has been found for the least squares norm, the iterations are continued at that delay with the Chebyshev norm.

Apart from saving a lot of time, first using the least squares norm has one additional advantage: it improves the numerical conditioning of the Chebyshev approximation problems and thus, it increases the maximal IIR filter length that can be designed. The Ellacot-Williams method, which is used extensively to solve the Chebyshev approximation problem, might get into trouble if too many poles are located excessively close to the unit circle [Ellacott and Williams, 1976b]. This typically occurs if a too small delay is applied to the target. However, the least squares method does not suffer from that situation, and therefore the least squares method still returns reliable results at delays where the Chebyshev method would fail. After a stable least squares delay is found, the zero/pole configuration at that delay will show much less poles clustered very close to the unit circle, and thus the Chebyshev method will have far less numerical problems to design the filter [Vuerinckx, 1998].

5.2.2 Initial Delay Value

Before starting with the actual iterations, an initial delay increment τ_0 is applied to the target. Ideally, this delay increment should already be a rough approximation of the stable delay increment; but unfortunately this stable delay is dependent on so many parameters, that it was impossible to construct a rule that approximates it. Instead, the initial delay increment τ_0 is chosen such that the mean group delay $\hat{\tau}_g$ of the target is set at a sensible value. The group delay $\hat{\tau}_g$ [Oppenheim and Shafer, 1989] is defined by

$$\hat{\tau}_g = -\frac{\psi_T(\omega_r) - \psi_T(\omega_l)}{\omega_r - \omega_l} \tag{5.6}$$

with $\psi_T(\omega)$ the unwrapped phase of the target frequency response $T(\omega)$ and with ω_r (respectively ω_l) the highest (respectively lowest) angular frequency where the target magnitude response is different from zero. Satisfactory results were obtained with the following initial delay increment τ_0

$$\tau_0 = n_\beta - n_\alpha - \hat{\tau}_g. \tag{5.7}$$

5.2.3 Finding the Delay Increment

The algorithm was published first in [Vuerinckx et al., 1996]. Given the unstable filter $H_{l-1}(z^{-1}, \theta)$, resulting from the previous iteration $l - 1$, a stable version $\hat{H}_{l-1}(z^{-1}, \theta)$ of it is produced by mirroring all unstable poles of $H_{l-1}(z^{-1}, \theta)$. This operation leaves the magnitude response of the filter $H_{l-1}(z^{-1}, \theta)$ unchanged, but not its phase response. Therefore, the stable filter $\hat{H}_{l-1}(z^{-1}, \theta)$ is not useful as an approximation of the target frequency response $T(\omega)$. However, along with the phase response, the group delay is also modified by mirroring the unstable poles and because this difference in group delay is a measure for the influence on the delay of unstable poles, it is used as an indication for the delay increment. Unfortunately, the group delay difference between both filters $H_{l-1}(z^{-1}, \theta)$ and $\hat{H}_{l-1}(z^{-1}, \theta)$ is not a constant term, which is the only allowed modification on the target. Therefore, an average of the group delay difference must be used instead. Let

$\tau_{g,l-1}$ and $\hat{\tau}_{g,l-1}$ be the mean group delay of $H_{l-1}(z^{-1}, \theta)$ and $\hat{H}_{l-1}(z^{-1}, \theta)$, respectively, calculated according to (5.6). The proposed delay increment $\Delta\tau_l$ to be added to the target is then

$$\Delta\tau_l = \hat{\tau}_{g,l-1} - \tau_{g,l-1}. \tag{5.8}$$

Practical examples showed that when getting in the vicinity of the stable delay, with all unstable poles very close to the unit circle, very small delay increments are generated by (5.8); resulting in a huge number of iterations before the stable delay is actually exceeded. In order to improve this behavior, equation (5.8) is corrected by imposing a lower bound on the increment, based on the length of the IIR filter. The final formula for calculating the delay increment then

$$\Delta\tau_l = \max\left\{\hat{\tau}_{g,l-1} - \tau_{g,l-1}, \frac{n_\beta + n_\alpha + 2}{100}\right\}. \tag{5.9}$$

Finally, the overall algorithm for finding a delay τ for which the Chebyshev IIR filter approximation becomes stable is given. But first of all a notation for the cost function defined with Chebyshev norm has to be introduced:

$$C_\infty(\theta, \tau) = \max_k |T(\Omega_k)e^{-j\omega_k\tau} - H(\Omega_k, \theta)|. \tag{5.10}$$

1. $\tau_0 = n_\beta - n_\alpha - \hat{\tau}_g$.

2. $\theta = \arg[\min_\theta C(\theta, \tau)]$.

3. Does θ determine a stable model? If yes then goto 8.

4. Determine $\hat{H}(\omega, \theta)$, τ_g, $\hat{\tau}_g$.

5. $\Delta\tau_l = \max\left\{\hat{\tau}_{g,l-1} - \tau_{g,l-1}, \frac{n_\beta+n_\alpha+2}{100}\right\}$.

6. $\tau = \tau + \Delta\tau$.

7. Goto 2.

8. $\theta = \arg[\min_\theta C_\infty(\theta, \tau)]$.

9. Does θ determine a stable model? If yes then goto 14.

10. Determine $\hat{H}(\omega, \theta)$, τ_g, $\hat{\tau}_g$.

11. $\Delta\tau_l = \max\left\{\hat{\tau}_{g,l-1} - \tau_{g,l-1}, \frac{n_\beta+n_\alpha+2}{100}\right\}$.

12. $\tau = \tau + \Delta\tau$.

13. Goto 8.

14. Stop.

5.3   Differential Equations

Here some theories about ordinary differential equations (ODE) are summarized. The complete theory of the ODE is very rich and extensive and because of the space limitation of this thesis only required slices are recalled. We need only the simplest differential equations: the so-called first order differential equations [Coddington and Levinson, 1995].

5.3.1   Problem Formulation

In the most general case an ordinary differential equation of the first-order contains an independent variable, an unknown function and its derivative and has the form [Hartman, 2002]

$$F(x, y(x), y'(x)) = 0, \tag{5.11}$$

where $F(x, y_0, y_1)$ is a given real valued function.

The results given here apply to the case where F is such that the equation $F(x, y_0, y_1) = 0$ can be solved for y_1 in the form $y_1 = f(x, y_0)$. Thus we shall consider differential equations of the form

$$y'(x) = f(x, y(x)). \tag{5.12}$$

Our goal is to find the unknown function $y(x)$ satisfying (5.12), i.e. we want to solve the following problem [Hartman, 2002], [Coddington and Levinson, 1995], [Zwillinger, 1997].

Problem 5.3.1. *Let I be a fixed interval of the real line and consider the following rectangle or strip on the plane*

$$D := \{(u, v) \in \mathbb{R}^2 | u \in I, v \in [c, d], -\infty \le c < d \le +\infty\}. \tag{5.13}$$

Suppose that f is a real valued function defined on D. Find a differentiable function y(x) defined on a real interval $J \subset I$ such that

- $(x, y(x)) \in D \ (x \in J)$,

- $y'(x) = f(x, y(x)) \ (x \in J)$.

This problem is called an *ordinary differential equation of the first-order*, and is usually written in the form (5.12). If such an interval J and a function exist, the $y(x)$ is called a *solution of the differential equation* (5.12) *on the interval J*. The graph of a solution of a differential equation is called an *integral curve* of the equation. If there are no such J and $y(x)$, we say that (5.12) has no solution [Hartman, 2002], [Coddington and Levinson, 1995], [Zwillinger, 1997].

Suppose that (τ, ξ) is a given point in D. Then an *initial value problem* associated with (5.12) and this point is defined in the following way.

Problem 5.3.2. *Find a solution y(x) of (5.12) satisfying the condition $y(\tau) = \xi$.*

The problem is denoted by

$$y'(x) = f(x, y(x)), \quad y(\tau) = \xi. \tag{5.14}$$

5.3.2 Existence and Uniqueness

The first question to be answered is under what conditions on f can we say that the problem has at least one solution.

The following theorem lays down a sufficient condition for a solution to exist [Hartman, 2002], [Coddington and Levinson, 1995], [Zwillinger, 1997].

Theorem 5.3.1 (Cauchy-Peano existence theorem). *If f is a continuous function on the strip D then there exists a solution of the initial value problem.*

The second question is the problem of uniqueness of the solutions. Not all problems possess an unique solution. The next theorems are about this question [Hartman, 2002], [Coddington and Levinson, 1995], [Zwillinger, 1997].

Definition 5.3.1. *Suppose f is defined on a strip D of the plane. If there exists a constant $L > 0$ such that for every (u, v_1) and (u, v_2) in D*

$$|f(u, v_1) - f(u, v_2)| \le L|v_1 - v_2| \tag{5.15}$$

then f is said to satisfy a Lipschitz condition (with respect to the second variable of f) in D. The constant L is called the Lipschitz constant.

The following fundamental existence and uniqueness theorem for the initial value problem given in (5.14) states that the problem (5.14) has exactly one solution, provided f satisfies a Lipschitz condition [Hartman, 2002], [Coddington and Levinson, 1995], [Zwillinger, 1997].

Theorem 5.3.2 (Picard-Lindelöf theorem). *Let f be a continuous function defined on the strip $D := \{(u, v) | a \le u \le b, v \in \mathbb{R}\}$, where a, b are finite real numbers. Suppose that f satisfies a Lipschitz condition on D. Then for every $\tau \in [a, b]$ and every $\xi \in \mathbb{R}$ there exists exactly one function $y(x)$ such that*

1. $y(x)$ is differentiable for $x \in [a, b]$,

2. $y'(x) = f(x, y(x))$ for $x \in [a, b]$,

3. $y(\tau) = \xi$.

Proposition 5.3.1. *If $f(x, y(x))$ and $\frac{\partial f(x,y)}{\partial y}$ are continuous on the domain D then $\forall (x_0, y_0) \in D$ problem (5.14) has exactly one solution on $[a, b]$.*

5.4   Numerical Methods of Differential Equations

Most of the differential equations have no explicit solution, i.e. there does not exist an analytical expression of $f(x)$. In these cases only a numerical solution can be computed which means that the true solution is approximated in some sense. There are many different methods that can be used to approximate solutions to a differential equation. Here we present two ones: the Euler method and the Runge-Kutta method. Both methods were realized in the developed algorithm presented in §5.6. All these discretize the differential system to produce a difference equation or map. The methods obtain different maps from the same differential equations, but they have the same aim: that the dynamics of the map should correspond closely to the dynamics of the differential equation [Cartwright and Piro, 1992].

Because of limitation in space the deep analysis of ODE solvers is not presented. It is important to note that through the introduction of the Euler method and the Runge-Kutta method, respectively, we assume that the conditions of existence and uniqueness are satisfied. In the case of the proposed algorithm it is verified in §5.6.2.

The notations and presented formulas are borrowed mainly from [Cartwright and Piro, 1992] and [Press et al., 2002].

5.4.1   Euler Method

Euler method is the oldest and the simplest method to produce a numerical solution of the initial value problem (5.14), [Kryszig, 1979]. From the calculus we know that the the tangent line of the solution $y(x)$ at x_0 can be formulated as

$$y = y_0 + y'(x_0)(x - x_0) \tag{5.16}$$

where $y_0 = y(x_0)$. And combining with (5.12) results

$$y = y_0 + f(x_0, y_0)(x - x_0). \tag{5.17}$$

The tangent is used to construct the estimation of $y(x)$ at the point $x_1 = x_0 + h$ where h is the step length. So, we compute the estimation of $y(x_1)$ as

$$y_1 = y_0 + f(x_0, y_0)(x_1 - x_0) = y_0 + f(x_0, y_0)h. \tag{5.18}$$

The schematic of the first of the Euler method can be seen in figure 5.2. The deduction of the first step (5.18) gives immediately the definition of every step of the Euler method:

$$y_{n+1} = y_n + f(x_n, y_n)h. \tag{5.19}$$

5.4.2   Runge-Kutta Method

The basic idea of the Runge-Kutta method is similar to that of the Euler method. In contrast with the Euler method the Runge-Kutta method uses not only one differentiation in order to estimate the solution $y(x)$ at the next point x_{n+1}.
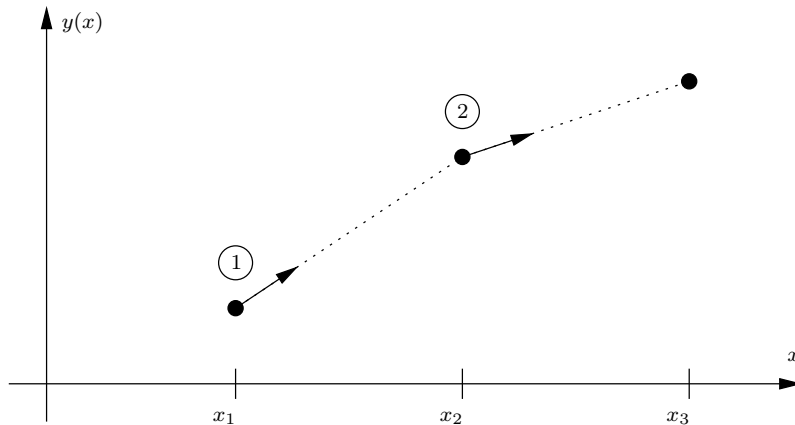
Figure 5.2: Illustration of the Euler method.

To obtain a q-stage Runge-Kutta method (q function evaluation per step) we let

$$Y_{n+1} = Y_n + h\phi(x_n, Y_n; h), \tag{5.20}$$

where

$$\phi(x_n, Y_n; h) = \sum_{i=1}^{q} v_i k_i, \tag{5.21}$$

so that

$$Y_{n+1} = Y_n + h\sum_{i=1}^{q} v_i k_i, \tag{5.22}$$

with

$$k_i = f\left(x_n + ha_i, Y_n + h\sum_{j=1}^{i-1} b_{ij}k_j\right) \tag{5.23}$$

and $a_1 = 0$ for an explicit method, or

$$k_i = f\left(x_n + ha_i, Y_n + h\sum_{j=1}^{q} b_{ij}k_j\right) \tag{5.24}$$

for an implicit method. For an explicit method, equation (5.23) can be solved for each k_i in turn, but for an implicit method equation (5.24) requires the solution of a non-linear system of k_is at each step. The set of explicit methods may be regarded as a subset of the set of implicit methods with $b_{ij} = 0$, $j \geq i$. In this thesis only explicit methods are developed and implemented. The Euler method is an explicit method, too.

An another example for an explicit method, the so-called midpoint method can be

obtained by letting $p = q = 2$. The illustration of the midpoint method can be seen in figure 5.3.
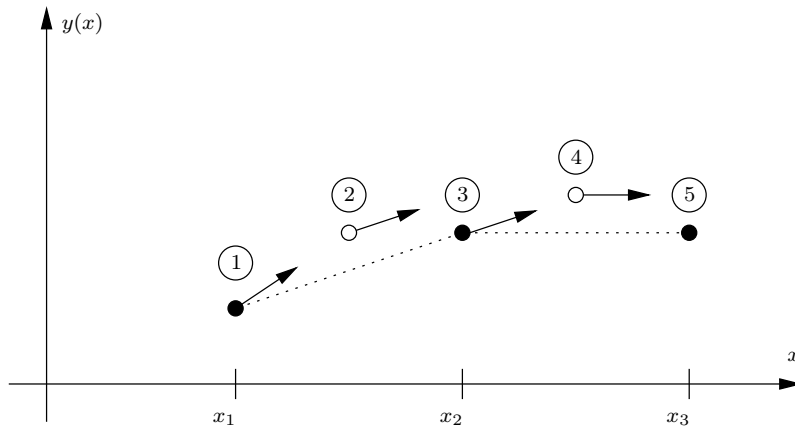
Figure 5.3: Illustration of the midpoint method.

For convenience, the coefficients a, b, and v of the Runge-Kutta method can be written in the form of a Butcher-array:

$$\begin{array}{c|c} a & b \\ \hline & v^T \end{array}$$

where $a = [a_1, a_2, \ldots, a_q]^T$, $v = [v_1, v_2, \ldots, v_q]^T$ and $b = [b_{ij}]$.

Runge-Kutta schemes are one-step and self-starting methods; they give Y_{n+1} in terms of Y_n only, and thus they produce a one-dimensional map if they are dealing with a single differential equation. The Runge-Kutta method is determined by its Bucher-array. To specifying the Butcher-array we have to solve a set of algebraic equations. Without going into details it has to be mentioned that increasing the number of stages of the Runge-Kutta methods increases its order which is a measurement of the accuracy of the solution. However, increasing the number of stages increasing the computing time and the difficulty of the corresponding algebraic equations. The fourth order Runge-Kutta method which is one of the algorithm used in the thesis is very popular because after that one has to add two more stages to the method to obtain any increase in the order.

In our algorithm the so-called Cash-Karps parameters are used. The Butcher-array is given in table 5.1.

5.5   Minimization of a Cost Function

As one can see in the previous subsection lots of practical algorithms in parametric identification and/or approximation are equivalent with a minimization of a so called cost function. In the approximation theory this cost function is a corresponding power of the norm. But in the parametric identification it is not so simple. In this thesis all the estimators can be formulated as the minimization of a cost function. The popular maximum

Table 5.1: Cash-Parameters for the Runge-Kutta method.

$\frac{1}{5}$	$\frac{1}{5}$				
$\frac{3}{10}$	$\frac{3}{40}$	$\frac{9}{40}$			
$\frac{3}{5}$	$\frac{3}{10}$	$-\frac{9}{10}$	$\frac{6}{5}$		
1	$-\frac{11}{54}$	$\frac{5}{2}$	$-\frac{70}{27}$	$\frac{35}{27}$	
$\frac{7}{8}$	$\frac{1631}{55296}$	$\frac{175}{512}$	$\frac{575}{13824}$	$\frac{44275}{110592}$	$\frac{253}{4096}$
	$\frac{37}{378}$	0	$\frac{250}{621}$	$\frac{125}{594}$	0	$\frac{512}{1771}$

likelihood estimator belongs to this set, too. In the identification part of the thesis more estimators with this property will be shown.

Some minimization problems lead to solving a linear matrix equation. The main advantage of this kind of estimators is that the solution can be computed efficiently and quickly. The general, so called least squares (LS) estimator exists in the frequency domain, too. The minimization of the LS cost function is equivalent to solving the matrix equation $Ap = b$ where A, b are given, p is the parameter vector. It is well known that the solution is $p = A^{+}b$ where $+$ in superscript denotes the Moore-Penrose pseudo inverse [Golub and Loan, 1996]. One of the generalizations of this problem is the total least squares problem (TLS) which is studied and generalized further in the identification part of the thesis. Unfortunately, most of the minimization problems cannot be formulated as a linear matrix equation; in these cases an iterative minimization has to be applied. In every step in the iteration a linear matrix equation is solved.

In this paper three commonly used minimization algorithms are applied. In the literature there are more proposed and fine tuned algorithms can be found in [Press et al., 2002]. It turned out that these three algorithms are very useful with respect to the frequency domain problems. These algorithms were tested in this context in the Frequency Domain Identification Toolbox [Kollár, 1994] which is used among others to identify linear systems.

The following three methods are studied below:

- Newton-Raphson,

- Newton-Gauss,

- and Levenberg-Marquardt.

The three methods minimize the same cost function and have similarities but different steps are done.

The three methods require a high quality initial parameter vector to produce good results. Without good initial value for the parameter vector, especially in the case of high order systems, there is less chance to find the global minima.

The methods presented here are described in [Schoukens and Pintelon, 1991] and [Pintelon and Schoukens, 2001].

5.5.1   Newton-Raphson

Let us consider a cost function $C(\theta)$, where $\theta \in \mathbb{R}^n$ is the parameter vector. Make the Taylor expansion at θ:

$$C(\theta + \Delta\theta) = C(\theta) + \left(\frac{\partial C(\theta)}{\partial \theta}\right)^T (\theta)\Delta\theta + \frac{1}{2}\Delta\theta^T \frac{\partial^2 C(\theta)}{\partial \theta^2}\Delta\theta + \cdots, \qquad (5.25)$$

where $\frac{\partial C(\theta)}{\partial \theta} \in \mathbb{R}^n$ and $\frac{\partial^2 C(\theta)}{\partial \theta^2} \in \mathbb{R}^{n \times n}$. The derivatives have to be taken with respect to θ. A necessary condition for having an extreme at $\theta + \Delta\theta$ is that the derivative with respect to $\Delta\theta$ at $\theta + \Delta\theta$ should be equal to zero. From the second order approximation of (5.25)

$$\frac{\partial C(\theta + \Delta\theta)}{\partial \Delta\theta} = 0 + \frac{\partial C(\theta)}{\partial \theta} + \frac{\partial^2 C(\theta)}{\partial \theta^2}\Delta\theta. \qquad (5.26)$$

The solution of this linear set of equations gives the value of $\Delta\theta$

$$\Delta\theta = -\left(\frac{\partial^2 C(\theta)}{\partial \theta^2}\right)^{-1} \frac{\partial C(\theta)}{\partial \theta}. \qquad (5.27)$$

The most simple Newton-Raphson algorithm consists in calculating successive values of $\Delta\theta$ (using subindex k to emphasis that it is a sequence):

1. selection of the starting values

2. calculation of $\Delta\theta_k$

3. $\theta_{k+1} = \theta_k + \Delta\theta_k$

4. if the stop condition not met, go back to 1

5. stop

5.5.2   Newton-Gauss

The Newton-Raphson algorithm has a major problem. Since the matrix $\left(\frac{\partial^2 C(\theta)}{\partial \theta^2}\right)$ might not be positive definite, the iterative algorithms may not converge. Solution to this problem is the Newton-Gauss algorithm. Consider the following form of the cost function:

$$C(\theta) = (u - g(\theta))^T (u - g(\theta)). \qquad (5.28)$$

If the vector $(u - g(\theta)) \in \mathbb{C}$ then the modified cost function is

$$C(\theta) = (u_C - g_C(\theta))^H (u_C - g_C(\theta)) \qquad (5.29)$$

where the superscript H denotes the Hermitee transformation (transposition and complex conjugation) of the complex matrix and C emphasizes that they are now complex variables.

This case can be reduced to the equation (5.28) if

$$u = \begin{bmatrix} \text{Real}\{u_C\} \\ \text{Imag}\{u_C\} \end{bmatrix} \quad g(\theta) = \begin{bmatrix} \text{Real}\{g_C(\theta)\} \\ \text{Imag}\{g_C(\theta)\} \end{bmatrix} \tag{5.30}$$

The second order derivatives are given by

$$\frac{\partial^2 C(\theta)}{\partial \theta^2} = 2 \sum_{k=1}^{N} \left(\left(\frac{\partial g_k(\theta)}{\partial \theta} \right)^T \left(\frac{\partial g_k(\theta)}{\partial \theta} \right) - (u_k - g_k(\theta)) \frac{\partial^2 g_k(\theta)}{\partial \theta^2} \right) \tag{5.31}$$

where N is the length of $g(\theta)$ and $g_j(\theta)$ is the jth element of $g(\theta)$. The second order derivatives can be approximated by

$$\frac{\partial^2 C(\theta)}{\partial \theta^2} \approx 2 \sum_{k=1}^{N} \left(\frac{\partial g_k(\theta)}{\partial \theta} \right)^T \left(\frac{\partial g_k(\theta)}{\partial \theta} \right)$$

$$= 2 \left(\frac{\partial g(\theta)}{\partial \theta} \right)^T \left(\frac{\partial g(\theta)}{\partial \theta} \right) = 2J(\theta)^T J(\theta) \tag{5.32}$$

with $J(\theta) = \frac{\partial g(\theta)}{\partial \theta}$ the Jacobian matrix. So the first order derivative of cost function is

$$\frac{\partial C(\theta)}{\partial \theta} = -2J(\theta)^T (u - g(\theta)) \tag{5.33}$$

and this results in the following modified step increase expression

$$\Delta \theta_k = (J_k^T J_k)^{-1} J_k^T (u - g(\theta_k)) = J_k^+ (u - g(\theta_k)) \tag{5.34}$$

where $J_k = J(\theta_k)$.

5.5.3 Levenberg-Marquardt

The iteration step is a combination of Newton-Gauss and the steepest descent. It is the following:

$$\Delta \theta_{k+1} = \frac{1}{\lambda} J_k^T (u - g(\theta_k)) = -\frac{1}{2\lambda} \text{grad}^T C(\theta_k) \tag{5.35}$$

hence

$$\Delta \theta_k = (J_k^T J_k + \lambda I)^{-1} J_k^T (u - g(\theta_k)) \tag{5.36}$$

where $\lambda > 0$ is a constant and I is eigenmatrix with appropriate size. If $\lambda \to 0$ then it is the Newton-Gauss method, and if $\lambda \to \infty$ it is the steepest descent. So the basic structure of a Levenberg-Marquardt algorithm is in general quite simple

1. selection of a set of starting values and a large starting value of λ

2. Calculate the Jacobian J_k

3. $\Delta\theta_{k+1} = (J_k^T J_k + \lambda I)^{-1} J_k^T (u - g(\theta_k))$

4. Is the step successful, i.e. $C(\theta_k) > C(\theta_{k+1})$?
   Yes : decrease λ and proceed to 5.
   No : increase λ and go to 3.

5. If it stop criterion is not met, go to 2.

6. Stop.

This method may have larger convergence domain than the previous ones.

5.6   Proposed Algorithm for Automatic Delay Selection

In this section the proposed algorithm is discussed. First, the deduction of the method is shown. After that, some theoretical establishments are given. And at the end of the section, as summary the flow-chart of the method is presented.

5.6.1   Deduction of the Method

Our aim is to solve the problem (5.3). In §5.2 a method based on the group delay is presented. Here, a new algorithm based on the ODE methods is developed and analyzed.

   The algorithm makes steps which are equivalent to a sequence of the parameter vectors θ and a sequence of the delays τ. As in [Vuerinckx, 1998], two kinds of steps are distinguished. Let $\theta_i(\tau)$ be the estimated parameter vector in the ith step.

- The first set of steps contains steps which try to minimize the cost function (5.1) without modifying the delay τ, i.e. we calculate $\theta_{i+1}(\tau)$ by using $\theta_i(\tau)$. These steps correspond to the methods presented in §5.5. Throughout this thesis these steps are called gradient steps because the gradient method serves as their basis.

- In steps from the second set the delay is incremented and we try to estimate $\theta_{i+1}(\tau + \Delta\tau)$ by using $\theta_i(\tau)$. We call these steps ODE steps because the developed method is practically one of the ODE solvers presented in §5.4.

It is worth noting that $\theta_i(\tau)$ can be the parameter vector where the cost function (5.1) is minimal ($\theta_i(\tau) = \theta_*(\tau)$) depending on the type of the step.

   And now we develop a method for steps from the second set. Equation (5.1) is continuous with respect to τ. The characterization of the extreme values of the cost function (with fixed τ) can be done by evaluating the following equation:

$$\frac{\partial C(\theta, \tau)}{\partial \theta} = 0. \tag{5.37}$$

This is a necessary but not a sufficient condition for a global minimum of the cost function. As we can see this equation defines an implicit function $\theta(\tau)$ which is continuous w.r.t.

the delay. It means that we can write a differential equation for the local minima. The following formulas stem from the total differentiation of equation (5.37):

$$\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2} \frac{d\theta}{d\tau} + \frac{\partial^2 C(\theta, \tau)}{\partial \theta \partial \tau} = 0. \tag{5.38}$$

Hence

$$\frac{d\theta}{d\tau} = - \left[\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2} \right]^{-1} \frac{\partial^2 C(\theta, \tau)}{\partial \theta \partial \tau} \tag{5.39}$$

or, if $\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2}$ is a singular matrix, then

$$\frac{d\theta}{d\tau} = - \left[\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2} \right]^{+} \frac{\partial^2 C(\theta, \tau)}{\partial \theta \partial \tau} \tag{5.40}$$

where the symbol + denotes the Moore-Penrose pseudo-inverse of a matrix.

This is a set of ordinary differential equations (ODE), and, hence numerical tools from the numerical differential equations theory can be used, [Press et al., 2002].

So, the overall algorithm is a sequence of two kinds of steps. If we find that the global minimum of the cost function determines an unstable model then the delay has to be increased and ODE steps have to be performed. Of course, the calculated parameter vector in an ODE step is only an estimation of the corresponding local minimum of the cost function. Since at every delay value the global minimum of the cost function can be calculated by using gradient steps, we have an advantage over the classical ODE solvers that in every x (in our case τ) the solution $y(x)$ (in our case $\theta_*(\tau)$) can be calculated without error. In the full search algorithm at every delay value τ the overall minimization process is evaluated. The full search is very time consuming and inserting ODE steps can reduce the computing time drastically.

Using (5.1) and the fact that

$$H(\Omega, \theta) = \frac{\beta_0 + \beta_1 \Omega + \ldots + \beta_{n_\beta} \Omega^{n_\beta}}{\alpha_0 + \alpha_1 \Omega + \ldots + \alpha_{n_\alpha} \Omega^{n_\alpha}} \tag{5.41}$$

the expression $\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2}$ and $\frac{\partial^2 C(\theta, \tau)}{\partial \theta \partial \tau}$ are calculated.

Equation (5.1) can be written as

$$C(\theta, \tau) = (u(\tau) - g(\theta))^T (u(\tau) - g(\theta)) \tag{5.42}$$

where $u(\tau) : \mathbb{R} \mapsto \mathbb{R}^{2F}$ and $g(\theta) : \mathbb{R}^{n_\beta + n_\alpha + 1} \mapsto \mathbb{R}^{2F}$. In this expression the usual map between the complex plane and the two dimensional real space is used. Differentiation of equation (5.42) w.r.t. θ gives

$$\left(\frac{\partial C(\theta, \tau)}{\partial \theta} \right)^T = -2 \left(\frac{\partial g(\theta)}{\partial \theta} \right)^T (u(\tau) - g(\theta)) \tag{5.43}$$

where $\frac{\partial g(\theta)}{\partial \theta} \in \mathbb{R}^{2F \times (n_\beta + n_\alpha + 1)}$. Then

$$\frac{\partial^2 C(\theta, \tau)}{\partial \theta^2} = -2 \sum_{k=1}^{2F} \left(\left(\frac{\partial^2 g_k(\theta)}{\partial \theta^2} \right) (u(\tau) - g(\theta))_k \right)$$

$$+ 2 \left(\frac{\partial g(\theta)}{\partial \theta} \right)^T \left(\frac{\partial g(\theta)}{\partial \theta} \right) \tag{5.44}$$

where the subscript k denotes the kth element of a vector and $\frac{\partial^2 g_k(\theta)}{\partial \theta^2} \in \mathbb{R}^{(n_\beta + n_\alpha + 1) \times (n_\beta + n_\alpha + 1)}$ and

$$\frac{\partial^2 C(\theta, \tau)}{\partial \theta \partial \tau} = -2 \left(\frac{\partial g(\theta)}{\partial \theta} \right)^T \frac{\partial u(\tau)}{\partial \tau}. \tag{5.45}$$

The first factor in (5.43) can be written as

$$\frac{\partial g_k(\theta)}{\partial \theta} = \frac{\partial H(\Omega_k, \theta)}{\partial \theta}. \tag{5.46}$$

Two cases have to be distinguished:

- Differentiating w.r.t. β_k and continuing (5.46) results

$$\frac{\partial H(\Omega, \theta)}{\partial \beta_k} = \frac{\Omega^k}{\alpha_0 + \alpha_1 \Omega + \ldots + \alpha_{n_\alpha} \Omega^{n_\alpha}}. \tag{5.47}$$

- Differentiating w.r.t. α_k results

$$\frac{\partial H(\Omega, \theta)}{\partial \alpha_k} = -H(\Omega, \theta) \frac{\Omega^k}{\alpha_0 + \alpha_1 \Omega + \ldots + \alpha_{n_\alpha} \Omega^{n_\alpha}}. \tag{5.48}$$

The second factor of (5.45) is

$$\frac{\partial y(\tau)}{\partial \tau} = \frac{\partial (T(\Omega) e^{-j\omega \tau})}{\partial \tau} = -j\omega T(\Omega) e^{-j\omega \tau}. \tag{5.49}$$

And in the end the first term in (5.44) has to be analyzed:

$$\frac{\partial^2 g_k(\theta)}{\partial \theta^2} = \frac{\partial^2 H(\Omega_k, \theta)}{\partial \theta^2}. \tag{5.50}$$

There are three cases:

- Differentiating w.r.t. β_m and β_n.

$$\frac{\partial^2 H(\Omega_k, \theta)}{\partial \beta_m \partial \beta_n} = 0. \tag{5.51}$$

- Differentiating w.r.t. α_m and β_n.

$$\frac{\partial^2 H(\Omega_k, \theta)}{\partial \beta_m \partial \alpha_n} = -\frac{\Omega_k^{m+n}}{(\alpha_0 + \alpha_1 \Omega_k + \ldots + \alpha_{n_\alpha} \Omega_k^{n_\alpha})^2}. \tag{5.52}$$

- Differentiating w.r.t. α_m and α_n.

$$\frac{\partial^2 H(\Omega_k, \theta)}{\partial \alpha_m \partial \alpha_n} = 2H(\Omega_k, \theta) \frac{\Omega_k^{m+n}}{(\alpha_0 + \alpha_1 \Omega_k + \ldots + \alpha_{n_\alpha} \Omega_k^{n_\alpha})^2} \tag{5.53}$$

5.6.2 Some Theoretical Underpinning

Here, the theoretical background of the ODE solver is checked and the effect of the finite frequency grid is analyzed.

5.6.2.1 Existence and Uniqueness

The cost function is a scale invariant function, i.e.

$$C(\theta, \tau) = C(\lambda \theta, \tau). \tag{5.54}$$

where λ is a non-zero, real number. The parameter ambiguity is removed by constraining the parameters: $\theta_j = 1$ (one coefficient is fixed), $\|\theta\|_2 = 1$ (the Euclidean norm of θ is constrained to one) or $\theta_i \theta_j = \pm 1$ ($i \neq j$, the multiplication of two coefficients is fixed), and so on. In this thesis the constraint $\|\theta\|_2 = 1$ is used. Hence, the possible set of parameters is a closed and bounded (=compact) which simplifies the treatment of the existence and uniqueness. Such a regular compact set is constructed as follows. Let $\Theta \subset \mathbb{R}^{n_\beta + n_\alpha + 1}$ be the unit ball that is a compact parameter set. Define $\Theta_s \subset \Theta$ as the singular set of parameter values for which the cost function (5.1) does not exists or is infinite. Usually, the topological dimension of this singular set is smaller than the dimension of Θ. The regular set Θ_r are those parameters in Θ which are not in an ε-distance of the singular set Θ_s

$$\Theta_r = \Theta \backslash \{\theta \in \Theta, \|\theta - \theta_s\| < \varepsilon, \theta_s \in \Theta_s\} \tag{5.55}$$

Θ_r is compact (closed and bounded) by construction. Using the same reasoning, a regular compact set is constructed where the (higher order) derivatives of the cost function exist and are finite.

In §5.3.2 some important theorems about existence and uniqueness of a solution of an ODE are listed. The proposition 5.3.1 of Picard-Lindelöf theorem is applied. The equation (5.39) and (5.40) are used. Obviously, the operations used from which the expressions are constructed are continuous functions. Furthermore, $\frac{\partial f(x,y)}{\partial y}$ exists and is continuous because of the same reasons. Now, we can conclude that the solution of the initial value problem (5.40) exists and is unique. The argumentation in the last paragraph ensures that there is no strange situation when the equation (5.40) has to be computed.

5.6.2.2   Effect of the Discrete Frequency Grid

There is a big difference between the cost function (5.1) presented in this chapter and the cost function (3.26) mentioned in the existence proof. In (5.1) the integral is replaced by the sum because of the discrete frequency grid and therefore the existence proof for continuous frequency interval is not valid.

In the case of a discrete frequency grid the existence theorem 4.3.1 cannot guarantee that for all τ sufficiently large the global minimum of the cost function determines a stable model. Unfortunately, at this moment there is neither a proof of a similar existence theorem nor a counter-example. From a mathematical point of view, the following is not true for the discrete frequency grid: there exists a delay τ sufficiently large such that for all frequency point there exists an appropriate sliced complex plane (in order to define the complex logarithm) and a neighborhood (in the frequency interval) where the phase of the target function is monotonously decreasing with the defined complex logarithm. In the case of a complex function defined on a finite set a monotone phase is meaningless. In the existence proof the continuous frequency interval enables us to use the Residue theorem 4.2.13 that constitutes the basis of the proof.

Fortunately, a method that results in a stable model which is the minimum of the cost function (5.1) in the desired frequency band can be given. Increasing the number of frequency points in the frequency grid such that the mesh size $\Delta\omega = \arg\max|\omega_{i+1} - \omega_i|$ decreases causes that the discrete cost function (5.1) converges to the cost function (3.26). In practical stable approximation applications the unstable model is available, hence the transfer function $T(\Omega)$ can be computed at arbitrary frequency points. Therefore, if for all τ the global minimum $\theta_*(\tau)$ determined an unstable model, the mesh size would decrease. Unfortunately, the examples that were investigated by the author did not have this property, so there is no practical observation of efficiency of this idea.

5.6.2.3   Causality

Although $H(\Omega, \theta)$ is a causal system by its structure, it may happen that in a step of the proposed method $\tau < 0$. $H(\Omega, \theta)$ approximates the causal part of $T(\Omega)e^{-j\omega\tau}$, hence the norm of the non-causal part is a lower bound of the approximation error. The approximation problem has a causal solution even if $\tau < 0$. In a particular step of the proposed solver it is possible that the delay decreases because the direction of the steps depends on the derivative of $C(\theta, \tau)$ at the point $[\theta, \tau]$. Moreover, in the case of a tricky $T(\Omega)$ it may happen that the solution of the optimization problem (5.3) leads to a negative delay value which is probably very close to zero.

Practical experience shows that decreasing the delay causes the solution of the optimization problem be an unstable model. Therefore, it is unlikely that in the case of real stable approximation problems a negative delay value is obtained as a solution.

5.6.3   Flow-chart of the Method

The algorithm can be divided into two parts. First, a stable model which is the global minimum of the function at a fixed delay value has to be found. After that, the solution

found may be improved by an extended gradient method. In the second step the extended gradient method means that the delay is added to the free parameter set and a global minimization procedure is performed with this new parameter vector. Of course, there is no guarantee that the parameter vector where the cost function is minimal determines a stable model. Therefore, a restricted method is used likewise in §3.2.3. An extended gradient step cannot increase the value of the cost function, hence the approximation error may be decreased but it is impossible to be increased.

The flow-chart of the first part is the following:

1. Start.

2. $\tau = 0$.

3. Calculate an initial value of the parameter vector $\theta_0(\tau)$.

4. Decide if it is an ODE step or gradient. (For example, every kth step, or/and if the variation of the cost function exceed a threshold.)

5. If it is an ODE step then goto 8.

6. Gradient steps until the minimum is reached.

7. Goto 9.

8. Make an ODE estimation.

9. Meet the stop condition? If not then goto 4.

10. Stop.

In step 4 the algorithm has to be decide about the kind of the next step. In the presented ODE solvers in every step (except the initial steps) values of the estimated solution are used. In this proposed method, at arbitrary fixed delay value it is possible to perform a complete gradient-based minimization method. Thus, at arbitrary delay value the *exact* solution of the problem 5.14 can be computed. Using only this method to find the solution is very slow. Less computation time is required in computing an ODE step, therefore the proposed algorithm contains both ODE and gradient steps. Furthermore, in the neighborhood of a parameter vector that determines an unstable model are parameter vectors to which unstable models belong and if not then a more profound search can be done.

5.7 Examples

To illustrate the competence of the algorithm in this section some examples are shown. Every example has been already published [Pintelon et al., 1990], [Kollár et al., 1991], [Kollár et al., 1990], [Pintelon and Schoukens, 1990], so we can make a comparison between the published results and the output of the proposed algorithm. This section contains three subsections which belong to different examples. In the first two subsections ideal

(unstable) transfer functions are approximated, and in the last one an example based on real measurements is introduced.

The relative complex error of the realized transfer function H is defined as

$$\delta_C = \frac{H - T}{T}. \tag{5.56}$$

The magnitude of the relative complex error is

$$\delta = |\delta_C| = \left| \frac{H - T}{T} \right|. \tag{5.57}$$

The latter quantity is used to describe the goodness of the approximations.

The realization of a digital filter needs poles which are not too close to the unit circle. Distance between the set of poles and the unit circle is defined as

$$D_{\mathbb{T}} = \min_{k=1,\dots,n_\alpha} ||p_k| - 1|. \tag{5.58}$$

In practice the necessary lower bound of $D_{\mathbb{T}}$ depends on the realization method. A possible choice is the cascade of second order sections. In this case a fixed-point DSP processor which is capable to perform 16-bit fractional multiplication, for realizable systems a good lower bound of $D_{\mathbb{T}}$ is 0.0001.

5.7.1   Hilbert Transformer

The Hilbert transform is used in communication and measurement applications. The standard method of the Hilbert transformer design is the Reméz algorithm, which is well elaborated. There are attempts to design IIR realizations, too [Ansari, 1987]. Chen and Parks [Chen and Parks, 1987] have shown that the performance of FIR filters can be improved by accomplishing the fit in the complex domain, and by allowing small deviations form the linear phase. This leads to the idea of designing both FIR and IIR Hilbert transformers by fitting the filter to the complex frequency response [Kollár et al., 1990]. In this thesis only the IIR case is investigated because FIR filters are always stable.

The design of a (two-sided) digital Hilbert transformer means that we try to approximate the following transfer function:

$$T(f) = \begin{cases} -j \text{ if } 0 < f < f_s/2 \\ j \text{ if } f_s/2 < f < f_s \end{cases} \tag{5.59}$$

where f_s is the sampling frequency. The magnitude characteristic of a band-pass Hilbert transformer can be seen in figure 5.4.

Because of the discontinuity at 0 and $f_s/2$, a digital filter can approximate this well only in a given band. Fortunately, in most applications this can be tolerated. In our example the frequency band is [0.04 : 0.005 : 0.46].
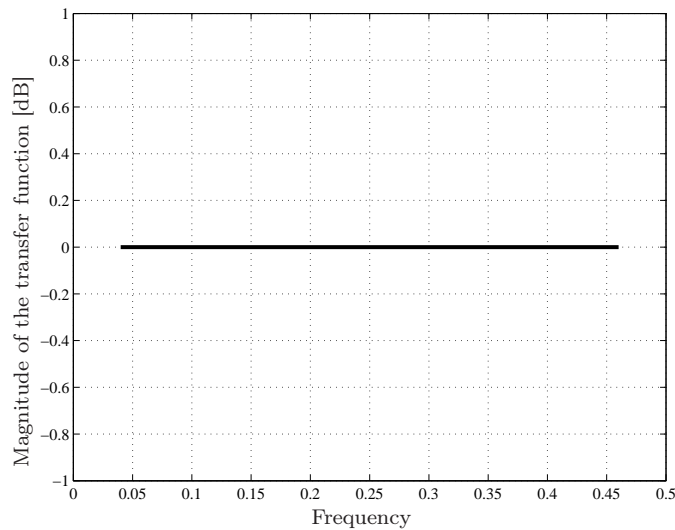
Figure 5.4: Magnitude characteristic of a band-pass Hilbert transformer. $f_s = 1$.

Table 5.2: Coefficients of the Hilbert transformer in the case $\tau = 10.7762$. $D_{\mathbb{T}} = 0.0523$.

number	zero	pole
1	$0.8024 \pm j1.5276$	0.8900
2	$1.3727 \pm j0.8379$	-0.9477
3	1.1235	$0.5307 \pm j0.3239$
4	$-0.0886 \pm j1.7653$	$0.2695 \pm j0.5130$
5	$-0.9602 \pm j1.4682$	$-0.5431 \pm j0.2714$
6	$-1.4731 \pm j0.7362$	$-0.3120 \pm j0.4770$
7	-1.0551	$-0.0284 \pm j0.5650$

In [Kollár et al., 1990] the given pass-band Hilbert transformer was approximated by a 12th order model. In this example the same orders were used. The automatic delay search started at value $\tau = 0$.

First, we study the case where we allow a fractional delay. The proposed algorithm finds a parameter vector that belongs to a stable model at value $\tau = 10.7762$. The relative complex error is shown in figure 5.5. This result corresponds to the result of [Kollár et al., 1990]. However, it was produced by a full automatic algorithm. The published results of [Kollár et al., 1990] were obtained by using an initial delay value set. The cost function was minimized in cases of all the delay values from this set. The delay was free during the minimizations. The roots of the numerator and the denominator are summarized in table 5.2.

Due to the fractional delay, the function in figure 5.5 is asymmetric. In lots of time when we use Hilbert transformers an integer delay value is necessary. The transfer function (5.59) can be approximated by a stable model at the delay value $\tau = 11$. The magnitude

Figure 5.5: Magnitude of the relative complex error of the Hilbert transformer in the case of fractional delay. $\tau = 10.7762$.

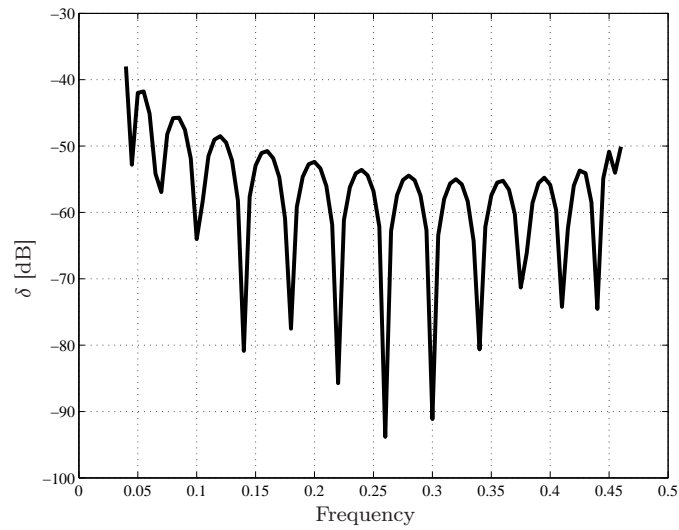of the relative complex error can be seen in figure 5.6 and roots of the numerator and the denominator are summarized in table 5.3.
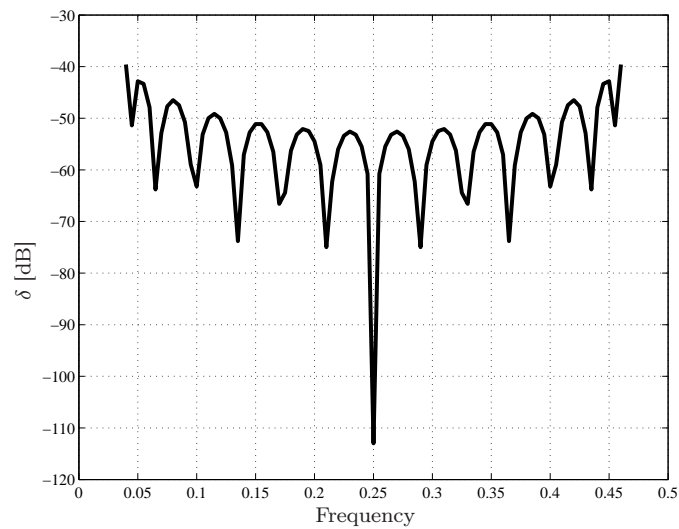

Figure 5.6: Magnitude of the relative complex error of the Hilbert transformer in the case of integer delay. $\tau = 11$.

In both cases it can be seen that roots of the denominators are not too close to the unit circle. Therefore the model can be realized in practise.

Table 5.3: Coefficients of the Hilbert transformer in the case $\tau = 11.0$. $D_\mathbb{T} = 0.1077$.

number	zero	pole
1	$-0.8399 \pm j1.4697$	-0.8923
2	$-1.3698 \pm j0.8031$	0.8923
3	-1.1206	$-0.5433 \pm j0.3185$
4	$-0.0000 \pm j1.7196$	$0.5433 \pm j0.3185$
5	$0.8399 \pm j1.4697$	$-0.2931 \pm j0.5129$
6	$1.3698 \pm j0.8031$	$0.2931 \pm j0.5129$
7	1.1206	$0.0000 \pm j0.5815$

5.7.2 Differentiators and Integrators

One is often not only interested in the physical quantity measured directly by the sensor, but also in its higher order derivative(s) or/and integral(s). An example is the study of the dynamic behavior of vibrating mechanical structures. Often the acceleration is measured but one is interested in the displacement of the structures: therefore the electronic signal supplied by the accelerometer is amplified and integrated twice. In a classical approach these operations are done by analog circuits, which involves problems with dc offset, noise amplification, accuracy, etc. Since the analog signal is digitized in the acquisition channel of the measurement device, one can think of performing the differentiation and the integration operation on the digitized signal. These methods are in fact nothing other than digital filter design. A possible solution is published in [Pintelon and Schoukens, 1990] and analyzed in this thesis, too.

5.7.2.1 Integrator

The basic idea is the following: the known ideal integrator is approximated only in a frequency band. The transfer function of the ideal integrator is:

$$T_{\text{integrator}}(\omega) = \frac{1}{j\omega}. \tag{5.60}$$

$T_{\text{integrator}}(\omega)$ is approximated only in the frequency band $[0.00125 : 0.00125 : 0.25]$. In [Pintelon and Schoukens, 1990] it was found that this integrator can be approximated well by using a $n_\beta = 6/n_\alpha = 6$ model and $\tau = 4.47796$ in the given frequency band.

An integrator always contains a pole at dc ($z = 0$). This may cause numerical overflow in the digital filter if a significant dc offset during a long time period is present. So, an additional high-pass filter must precede the digital integrator. This can be a linear phase FIR filter available in the literature [Rabiner and Gold, 1975] or an IIR filter designed by the proposed algorithm. Here, to avoid singularity at $\omega = 0$ and to increase numerical stability the target $T_{\text{integrator}}(\omega)$ is replaced by

$$\widetilde{T}_{\text{integrator}}(\omega) = \frac{1 - e^{j\omega}}{j\omega} \tag{5.61}$$

Table 5.4: Coefficients of the half-band integrator for $\tau = 4.47796$. $D_{\mathbb{T}} = 0.8743$.

number	zero	pole
1	$6.7983 \pm j4.1899$	$-0.0506 \pm j0.1151$
2	$-0.8069 \pm j6.4705$	$0.0799 \pm j0.0586$
3	-3.6641	

Table 5.5: Coefficients of the half-band integrator for $\tau = 2.5853$. $D_{\mathbb{T}} = 0.2180$.

number	zero	pole
1	$15.8825 \pm j22.3437$	-0.7820
2	-14.3795	-0.2140
3	-0.8694	$0.0158 \pm j0.0295$
4	-0.2968	

and an approximator $\widetilde{H}$ of order $n_\beta/(n_\alpha - 1)$ is calculated. The magnitude characteristic of $\widetilde{H}$ is in figure 5.7.
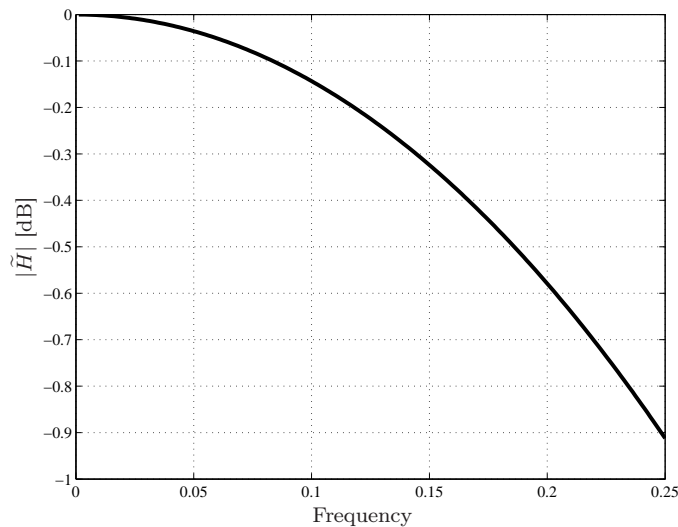
Figure 5.7: Magnitude of the half-band integrator $\widetilde{H}$.

The proposed algorithm finds a stable solution for $\tau = 2.5853$ that has a relative complex approximation error δ that is 20 dB smaller than the solution for $\tau = 4.47796$ given in [Pintelon and Schoukens, 1990]. The relative complex error can be seen in figure 5.8. Table 5.4 and table 5.5 summarize the poles and zeros of the resulting transfer functions. We can conclude that the advantage of the proposed algorithm is the automatic search of τ but it may be able to find a better solution than starting the minimization from a manually set-up delay value.
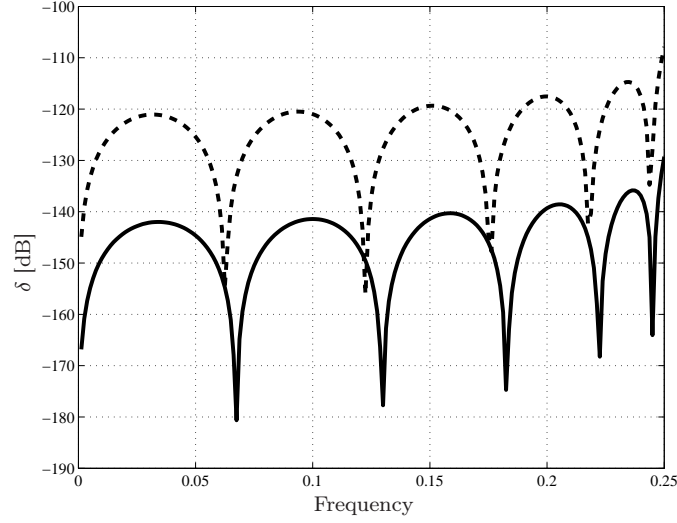
Figure 5.8: Magnitude of the complex relative errors of the half-band integrators. Dashed for $\tau = 4.47796$ and solid line for $\tau = 2.5853$.

5.7.2.2 Differentiator

The transfer function of an ideal differentiator is

$$T_{\text{differentiator}}(\omega) = j\omega. \tag{5.62}$$

As in the case of integrator $T_{\text{differentiator}}(\omega)$ is approximated in the frequency band $[0.00125 : 0.00125 : 0.25]$. The differentiator always contains a zero at $z = 1$. To decrease the dynamics of $T_{\text{differentiator}}(\omega)$, i.e. to increase the numerical stability of computations, like in the case of integrators, we uses this information and construct a modified target function:

$$\widetilde{T}_{\text{differentiator}}(\omega) = \frac{j\omega}{1 - e^{j\omega}} \tag{5.63}$$

and an approximator $\widetilde{H}$ of order $(n_\beta - 1)/n_\alpha$ is calculated. The magnitude characteristic of $\widetilde{H}$ is in figure 5.9.

In [Pintelon and Schoukens, 1990] a fifth order model was published with fractional delay $\tau = 3.41945$. Unfortunately, the proposed method cannot find this solution with using $\tau = 0$ as a starting value. The result of the algorithm is a stable model with $\tau = 2.38693$. The relative complex errors are in figure 5.10. The reason is that there is no guarantee that the local optimum with the lowest additional delay value is simultaneously a global optimum. Table 5.6 and table 5.7 summarize the poles and zeros of the resulting transfer functions.

It is interesting that if we use $\tau = 2.5$ as a starting value in the proposed algorithm then the solution for $\tau = 3.41945$ is found. It means that the next stable interval of domain of the function $\theta_*(\tau)$ contains the known best approximation.
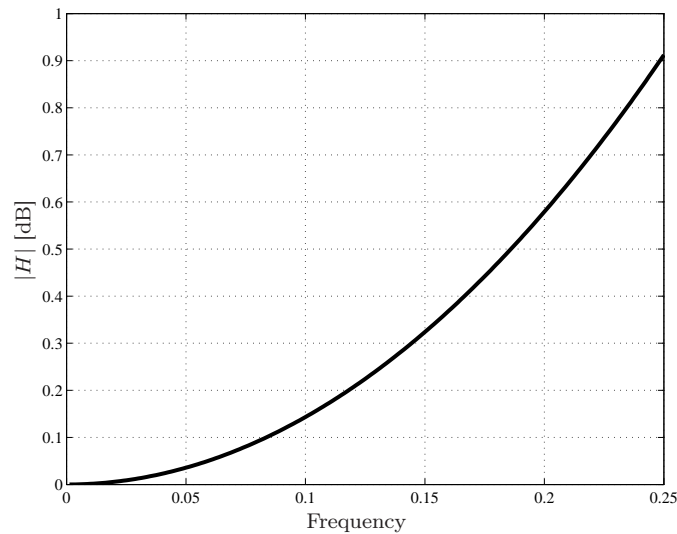
Figure 5.9: Magnitude of the half-band differentiator $\widetilde{H}$.
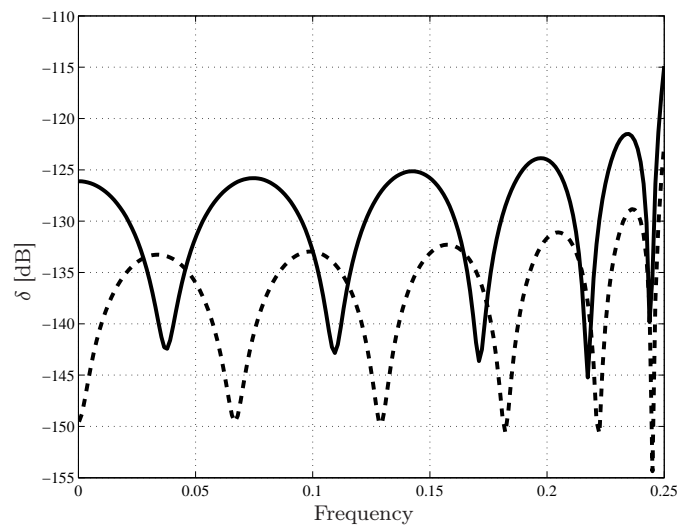
Figure 5.10: Magnitude of the complex relative errors of the half-band differentiators. Dashed for $\tau = 3.41945$ and solid line for $\tau = 2.38693$.

5.7.3 Data Acquisition Channels

This part handles the equalization of a data acquisition channel (DAC). For the first time, the idea, the measurement and the identification results were published in [Pintelon et al., 1990]. Some modifications can be found in [Kollár et al., 1991]. Moreover, there is an additional publication related to this problem, [Kollár et al., 1990]. Now, we summarize the result of mentioned publications and after that results of the proposed algorithm is shown. In §3.1.1 compensation or equalization of a channel is already mentioned as an application of the stable approximation. This subsection is devoted to illustrate the application of the

Table 5.6: Coefficients of the half-band differentiator for $\tau = 3.41945$. $D_{\mathbb{T}} = 0.2369$.

number	zero	pole
1	21.0455	-0.7631
2	$-2.9339 \pm j17.2560$	-0.1894
3	-0.6697	$0.0119 \pm j0.0826$
4		0.0703

Table 5.7: Coefficients of the half-band differentiator for $\tau = 2.38693$. $D_{\mathbb{T}} = 0.0597$.

number	zero	pole
1	$11.6417 \pm j34.3272$	-0.9403
2	-0.8135	-0.2403
3	-0.0475	$0.0004 \pm j0.0936$
4		0.0699

proposed algorithm in this field.

In [Pintelon et al., 1990] and [Kollár et al., 1990] compensation of an audio band Cauer filter of order 11, produced with laser trimmed thick film technology was presented. In [Pintelon et al., 1990] an amplitude- and phase-compensating IIR filter was designed. The relative complex error has been reduced to be less than -58 dB using a 14/14 IIR filter for the amplitude compensation and a 20/20 one for the phase equalization. Thus the resulting compensation filter has the order 34/34. It is obvious that this very straightforward method is not globally optimal, since the two separate steps have different goals. This inspired the article [Kollár et al., 1991]. In [Kollár et al., 1990] the order of the compensation filter was decreased by designing a new IIR filter, which compensated both the amplitude and the phase, with a lower order than 34/34. A 26/26 filter was obtained with practically the same fitting error as the 34/34 one.

In this thesis the design of the compensation filter was done by the presented algorithm. Before the stable approximation step an identification process has to be evaluated in order to achieve the optimal noise removal (see §3.1.2). The identification was done by the Frequency Identification Toolbox. The result is an IIR filter of order 14/14. The identified transfer function can be seen in figure 5.11 and the identified model parameters are summarized in table 5.8. The frequency grid is $[400 : 400 : 19200]$ Hz where the sampling frequency is 51200 Hz.

The resulted IIR filter is a nonminimal phase one, its inverse is unstable. Therefore stable approximation is needed. The approximation was computed by the proposed algorithm. In [Kollár et al., 1991] a filter of order 26/26 is proposed and the absolute value of the complex approximation error is less then 6 mdB which means that the relative complex error is approximately -58 dB. The results can be seen in figure 5.12. In the case of integer delay values (dashed line) the approximation error is approximately the same as in [Kollár
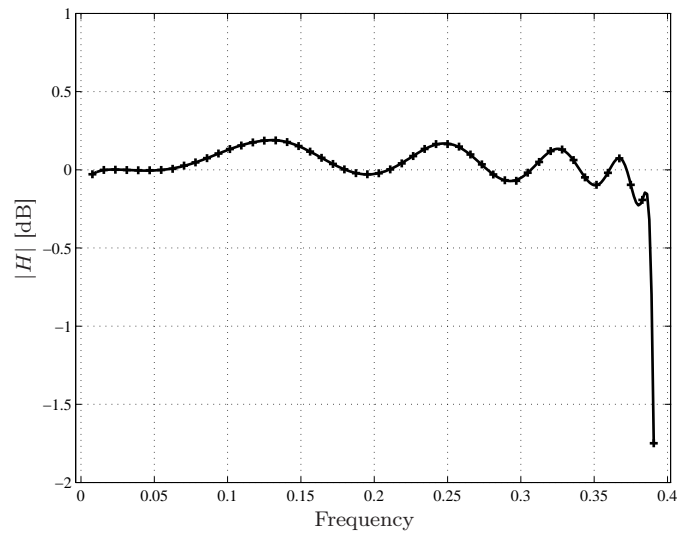
Figure 5.11: Estimation result of the identification step. The model is an IIR filter of order 14/14.

et al., 1991]. However, if a fractional delay is allowed then the algorithm returns a better approximation for which the magnitude of the relative complex error is less then -80 dB. The zeros and poles of the models are in table 5.9 and in table 5.10.

Table 5.8: Estimated zeros and poles of the DAC.

number	zero	pole
1	-28.0200	$-0.7425 \pm j0.6122$
2	-4.7072	$-0.6191 \pm j0.6319$
3	$-1.5442 \pm j0.6330$	$-0.3926 \pm j0.6720$
4	0.9473	0.9467
5	$0.6363 \pm j0.6578$	$0.6362 \pm j0.6578$
6	$-0.9347 \pm j0.4815$	$-0.0386 \pm j0.6582$
7	$-0.5963 \pm j0.4290$	0.4963
8	$-0.0941 \pm j0.3983$	$0.3330 \pm j0.4042$
9	0.3272	

Table 5.9: Zeros and poles of the data acquisition channel for $\tau = 35$. $D_{\mathbb{T}} = 0.1051$.

number	zero	pole
1	$-0.8470 \pm j0.8940$	-0.8949
2	$-0.6796 \pm j1.0667$	$-0.4995 \pm j0.6640$
3	$-0.4736 \pm j1.1898$	$-0.3075 \pm j0.7165$
4	$-0.2451 \pm j1.2662$	$-0.1760 \pm j0.7508$
5	$-0.0049 \pm j1.2958$	$-0.0211 \pm j0.7522$
6	$0.2371 \pm j1.2784$	$0.1321 \pm j0.7450$
7	$0.4718 \pm j1.2152$	$0.2723 \pm j0.7062$
8	$0.6909 \pm j1.1086$	$0.4042 \pm j0.6378$
9	$0.8868 \pm j0.9624$	$0.5235 \pm j0.5422$
10	$1.0529 \pm j0.7815$	$0.6210 \pm j0.4335$
11	$1.3503 \pm j0.1016$	$0.7590 \pm j0.1687$
12	$1.2800 \pm j0.3389$	0.7892
13	$1.1843 \pm j0.5717$	$0.7008 \pm j0.3103$
14		$0.0553 \pm j0.2704$

Table 5.10: Zeros and poles of the data acquisition channel for $\tau = 31.285868488$. $D_{\mathbb{T}} = 0.0579$.

| number	zero	pole
1	$-0.8472 \pm j0.9889$	$-0.5186 \pm j0.6421$
2	$-0.6436 \pm j1.1940$	$-0.5615 \pm j0.5439$
3	$-0.3916 \pm j1.3260$	-0.6014
4	$-0.1129 \pm j1.3929$	$-0.2691 \pm j0.6962$
5	$-0.5922 \pm j0.6065$	$-0.1453 \pm j0.7015$
6	$0.1758 \pm j1.3957$	$0.0126 \pm j0.6817$
7	$0.4598 \pm j1.3361$	$0.1407 \pm j0.6828$
8	$0.7255 \pm j1.2173$	$0.2685 \pm j0.6489$
9	$0.9609 \pm j1.0452$	$0.3985 \pm j0.5844$
10	$1.1553 \pm j0.8275$	$0.5258 \pm j0.4835$
11	$1.2997 \pm j0.5739$	$0.6172 \pm j0.3588$
12	$1.3875 \pm j0.2946$	0.9421
13	1.4162	$0.6772 \pm j0.2266$
14	0.9414	$0.7138 \pm j0.0809$
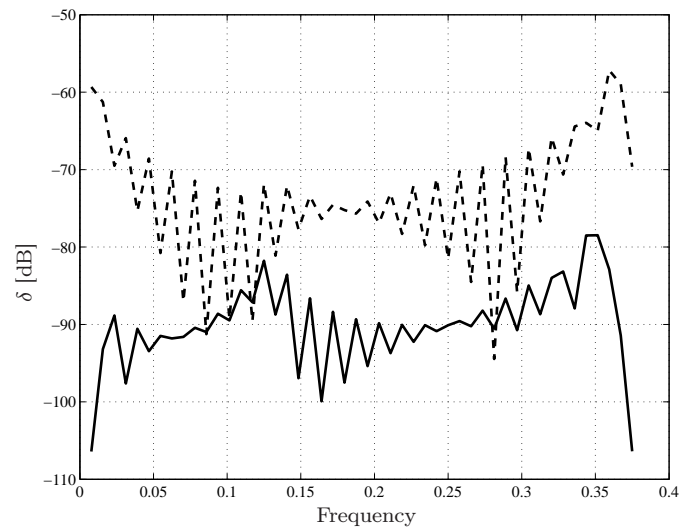
Figure 5.12: Magnitude of the complex relative errors of the DAC compensation. Dashed for $\tau = 35$ and solid line for $\tau = 31.285868488$.

Chapter 6

Basis Transformation for the TLS Problem

6.1 Introduction of the TLS method

The total least squares (TLS) problem can be discussed as a general statistical method to solve over-determined noisy linear equations or as a special method of frequency domain identification techniques. Here, the latter one was chosen but some comments about the general framework are mentioned, too [Pintelon et al., 1998], [Van Huffel and Vandewalle, 1991].

The method can be formulated as the minimization of a cost function with two constraints. If the cost function is the Froubenius norm, the solution is calculated using the singular value decomposition (SVD). It means that the result of the estimator can be obtained by using a reliable numerical tool. In some cases the parameter vector space has to be transformed. Examples are frequency scaling, applying orthogonal polynomials, and the case of a known subsystem. The former two transformations are typically applied in order to increase numerical stability of the computation. They are linear, therefore they can be expressed as multiplication by a transformation matrix.

In the TLS problem the solution space is restricted to the unit sphere in the parameter space. This constraint removes the ambiguity from model space (5.54). If the parameter space transformed, then intuitively, the constraint should be transformed, too. If not, then the solution does not coincide with the solution of the original TLS problem. This chapter explains and determines the necessary transformation in a proper way so that the solution equals the solution of the original TLS problem.

First, this chapter contains a little overview about the frequency domain system identification. After that, the maximum likelihood and the TLS estimators are introduced. The next part is about the parameter transformation, i.e. when and in which problems it is used. In the end the proposed correction is analyzed and examples are showed.

It is worth noting that the topic of this chapter differs from the two previous ones, therefore sometimes different notations are used or some notations are redefined according to common notations in frequency domain system identification. Mostly, notations in [Van

Huffel and Vandewalle, 1991] and [Pintelon and Schoukens, 2001] are used in this part of the thesis.

6.2   Frequency Domain System Identification

The basis for the treated identification approach is the frequency domain that is introduced in details in chapter 2. This means that the computed or measured frequency response functions (FRF) are used throughout the algorithms. The basic aim is to measure and model the transfer function of a plant, starting from noisy input and output measurements, see figure 6.1. $H_0(\Omega)$ is called the device under test (DUT) and is the true transfer function. Our aim is to determine an estimated transfer function $H(\Omega, \theta)$ such that the difference between $H(\Omega, \theta)$ and $H_0(\Omega)$ should be minimal in some sense [Doob, 1953], [Feller, 1068].
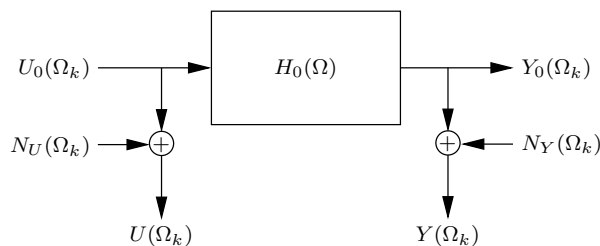

Figure 6.1: Frequency domain representation of the measurement process. $U_0(\Omega_k)$ and $Y_0(\Omega_k)$ are the true input and output, respectively. $U(\Omega_k)$, $Y(\Omega_k)$ are the measured quantities. $N_U(\Omega_k)$, $N_Y(\Omega_k)$ denote the noise processes. $H_0(\Omega)$ is the device under test.

Although the time domain identification techniques [Ljung, 1999] can still be used for periodic excitations under the zero-order-hold (ZOH) and band limited (BL) signal assumption, it is strongly recommended in this case to leave the time domain in favor of the frequency domain. Indeed in the frequency domain, one has the following features not found in the time domain [Pintelon et al., 1994]

- Since the non-excited frequency lines are eliminated, the noises are reduced [Ljung, 1993], [Schoukens et al., 1994].

- Saving frequency domain data reduces the necessary space because a large number of time-domain samples are replaced by a small number of spectral lines [Ljung, 1993], [Schoukens et al., 1994].

- When using a DFT to calculate the spectra, the frequency-domain noise is asymptotically complex normally distributed [Brillinger, 1981].

- Model validation: Using periodic excitations one has very good point estimates of the frequency response function [Schoukens et al., 1994], [Guillaume et al., 1992].

- It is very easy to combine data from different experiments [Ljung, 1993], [Schoukens et al., 1994].

Instead of using the direct division of both measured spectra $U(\Omega_k)$ and $Y(\Omega_k)$ the errors-in-variables approach is applied. The input and output spectra are considered as unknown parameters, connected by the parametric transfer function model:

$$Y(\Omega_k) = Y_0(\Omega_k) + N_Y(\Omega_k)$$

$$U(\Omega_k) = U_0(\Omega_k) + N_U(\Omega_k) \tag{6.1}$$

with $Y_0(\Omega_k) = H_0(\Omega_k)U_0(\Omega_k)$ and where $N_Y(\Omega_k)$ and $N_U(\Omega_k)$ include the generator noise, the process noise, and the measurement noise [Pintelon et al., 1994].

Without assumptions about the noise process, statistical properties of estimators cannot be studied and compared. In this case of time domain measurements we have the following assumptions: at the sampling instances the distributing time domain noise source $n_y(t)$, $n_u(t)$ are jointly correlated, and modelled as filtered white noise sequence

$$\begin{bmatrix} n_y(t) \\ n_u(t) \end{bmatrix} = \begin{bmatrix} H_{n,11}(z) & H_{n,12}(z) \\ H_{n,21}(z) & H_{n,22}(z) \end{bmatrix} \begin{bmatrix} e_1(t) \\ e_2(t) \end{bmatrix} \tag{6.2}$$

with $e^T(t) = [e_1(t), e_2(t)]$ and where $H_n(z^{-1})$ is a stable filter. $e(t)$ is independently distributed (over t and over its entries) with continuous probability density function, has stationary first and second order moments, uniformly bounded fourth order moments, and is independent of the true (unknown) excitation $u_0(t)$ [Pintelon and Schoukens, 2001]. The relationships between the frequency domain errors and time domain noise are

$$N_Y(\Omega_k) = \text{DFT}(n_y(t))$$

$$N_U(\Omega_k) = \text{DFT}(n_u(t)). \tag{6.3}$$

In the case of frequency domain experiments we have the following assumptions: $N_U(\Omega_k)$, $N_Y(\Omega_k)$ are independent (over k), jointly correlated, zero mean random variables with uniformly bounded absolute moments of order four. $N_Y(\Omega_k)$, $N_U(\Omega_k)$ are independent of the true excitation $U_0(\Omega_k)$ [Pintelon and Schoukens, 2001].

The estimators are compared by using the so-called cost functions. Every treated estimator can be written as an optimization problem which is minimization in this case. However, some estimations are computed not by solving an optimization problem but by solving a linear equations. One of the advantage of the TLS estimator is that the estimated parameter vector can be computed by solving the corresponding linear equations.

Like in the previous chapters a cost function is nearly related to norms. In a stochastic framework we usually try to minimize the norm of the error which is a vector in a finite dimensional space. Here, the norm L_2 is used because of its good properties.

6.2.1 Maximum Likelihood Estimator

The maximum likelihood estimator is a general method in statistic. Under the previously enumerated assumptions one can construct the Gaussian likelihood function, which gives the probability of obtaining a set of measurements $\{U(\Omega_k), Y(\Omega_k), k = 1, \ldots, F\}$, for a given set of values of the spectra $\{U_0(\Omega), Y_0(\Omega), k = 1, \ldots, F\}$. The likelihood function is a function of the (unknown) input and output spectra $U(\Omega)$ and $Y(\Omega)$, and the model parameters through the model equation

$$H(\Omega, \theta) = \frac{Y(\Omega)}{U(\Omega)}. \tag{6.4}$$

Maximization of the likelihood function with respect to $U(\Omega)$ and $Y(\Omega)$ under the constraint (6.4) results in the following maximum likelihood (ML) cost function [Pintelon and Schoukens, 2001]

$$C_{\mathrm{ML}}(\theta) = = \sum_{k=1}^{F} \frac{|N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)|^2}{\sigma_{U,k}^2 |N(\Omega_k, \theta)|^2 + \sigma_{Y,k}^2 |D(\Omega_k, \theta)|^2 - 2\mathrm{Re}\{\rho_{YU,k}D(\Omega_k, \theta)N^*(\Omega_k, \theta)\}} \tag{6.5}$$

where

$$\sigma_{U,k}^2 = \mathbb{E}\{N_U^*(\Omega_k)N_U(\Omega_k)\}, \ \ \sigma_{Y,k}^2 = \mathbb{E}\{N_Y^*(\Omega_k)N_Y(\Omega_k)\},$$

$$\rho_{YU,k} = \mathbb{E}\{N_Y(\Omega_k)N_U^*(\Omega_k)\} \tag{6.6}$$

and $\mathbb{E}$ denotes the expected value of a random variable. It is worth noting that the denominator of the cost function (6.5) is the variance of the numerator. It is called the optimal noise weighting. Hence the contribution of the equation error at frequency Ω_k to the ML cost is entirely determined by its variance: much confidence is given to accurate measurements while noisy measurements are rejected [Pintelon and Schoukens, 2001].

There is a disadvantage of the ML estimator: the cost function (6.5) is non-linear in the parameters θ. One of the algorithms presented in §5.5 can be used to compute the estimation numerically.

6.2.2 Total Least Squares Estimator

The expression *total least squares* is used not only in the theory of frequency domain identification but in the general theory of linear equations. The frequency domain method is an element of the general theory. This subsection begins with the general theory that contains the LS and TLS methods. After that, the application for frequency domain identification is shown.

6.2.2.1   Singular Value Decomposition

The singular value decomposition (SVD) is a very useful method which helps with characterizing matrices and solving linear equations.

Definition 6.2.1 (Singular value decomposition). *If $A \in \mathbb{R}^{m \times n}$ $(m > n)$ then the decomposition*

$$A = U \Sigma V^T \tag{6.7}$$

with

$$U = [U_1; U_2], \; U_1 = [u_1, \ldots, u_n], \; U_2 = [u_{n+1}, \ldots, u_m], \; u_i \in \mathbb{R}^m, \; U^T U = I_m$$

$$V = [v_1, \ldots, v_n], \; v_i \in \mathbb{R}^n, \; V^T V = I_n$$

$$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{m \times n}, \; \sigma_1 \geq \cdots \geq \sigma_n \geq 0. \tag{6.8}$$

is called singular value decomposition. σ_i are the singular values of A. The vector u_i is the ith left singular vector, and the vector v_i is the ith right singular vector [Van Huffel and Vandewalle, 1991], [Golub and Loan, 1996].

It is worth noting that the SVD exists for every matrix. Two example which show how it is possible to determine some characteristics of the matrix from the SVD:

- The rank of the matrix A equals with the number of the non-zero elements of Σ.

- The pseudo inverse [Golub and Loan, 1996] of the matrix A may be computed by SVD:

$$A^+ = V \mathrm{diag}(1/\sigma_1, \ldots, 1/\sigma_l, 0, \ldots, 0) U^T \tag{6.9}$$

  where l is the number of the non-zero singular values of A.

Definition 6.2.2 (Generalized singular value decomposition). *If we have $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ and $B \in \mathbb{R}^{p \times n}$, then there exist orthogonal $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{p \times p}$ and an invertible $X \in \mathbb{R}^{n \times n}$ such that*

$$U^T A X = \mathrm{diag}(c_1, \ldots, c_n) \quad c_i \geq 0 \tag{6.10}$$

and

$$V^T B X = \mathrm{diag}(s_1, \ldots, s_n) \quad s_i \geq 0 \tag{6.11}$$

where $q = \min(p, n)$ [Golub and Loan, 1996].

6.2.2.2   TLS Method of Over-determined Linear Equations

A basic problem of applied mathematics is to determine an estimate of the true but unknown parameters $x \in \mathbb{R}^n$ from certain measurements of the variables. This gives rise to an over-determined set of m linear equations $(m > n)$:

$$Ax \approx b \tag{6.12}$$

where the matrix $A \in \mathbb{R}^{m \times n}$ and the vector $b \in \mathbb{R}^m$. In the classical LS approach the measurement A are assumed to be free of error; hence all errors are confined to the observation vector b. However, this assumption is frequently unrealistic: sampling errors, human errors, modelling error, and instrument errors may imply inaccuracies of the data matrix A as well. TLS is a method of fitting that is appropriate when there are errors in both the observation vector b and the data matrix A.

In this section we formulate the main principle of TLS method. A good way to introduce the method is to recast the ordinary least squares problem.

Problem 6.2.1 (Ordinary least squares problem). *Given an over-determined set of m linear equations $Ax \approx b$ in n unknowns x, the least squares problem seeks to*

$$\underset{b' \in \mathbb{R}^m}{\text{minimize}} \|b - b'\|_2 \text{ subject to } b' \in \mathcal{R}(A). \tag{6.13}$$

where $\mathcal{R}(A)$ denotes the range of A. Once a minimizing b' is found, then any x satisfying

$$Ax = b' \tag{6.14}$$

is called LS solution and $\Delta b' = b - b'$ the corresponding LS correction [Van Huffel and Vandewalle, 1991].

Equation (6.13) is satisfied if b' is the orthogonal projection of b onto $\mathcal{R}(A)$. Thus, the LS problem amounts to perturbing the observation vector b by a minimum amount $\Delta b'$ so that $b' = b - \Delta b'$ can be "predicted" by the columns of A.

One way to take errors in A into account is to introduce perturbations in A also and to consider the following TLS problem.

Problem 6.2.2 (Basic total least squares problem). *Given an over-determined set of m linear equations $Ax \approx b$ in n unknowns x, the total least squares problem seeks to*

$$\underset{[\hat{A};\hat{b}] \in \mathbb{R}^{m \times (n+1)}}{\text{minimize}} \|[A;b] - [\hat{A};\hat{b}]\|_F \text{ subject to } \hat{b} \in \mathcal{R}(\hat{A}). \tag{6.15}$$

Once a minimizing $[\hat{A};\hat{b}]$ is found, then any x satisfying

$$\hat{A}x = \hat{b} \tag{6.16}$$

is called a TLS solution and $[\Delta\hat{A}; \Delta\hat{b}] = [A;b] - [\hat{A};\hat{b}]$ the corresponding TLS correction [Van Huffel and Vandewalle, 1991].

The TLS solution is denoted by $\hat{x}$. One important application of TLS problems is parameter estimation in errors-in-variables models. Here, we assume that m measurements in A, b are related to n unknown parameters x by

$$A_0 x = b, \quad A = A_0 + \Delta A, \quad b = b_0 + \Delta b \tag{6.17}$$

where ΔA, Δb represent the measurement errors. With some restrictions to the distribution of the errors the TLS estimates is consistent.

The next theorem is about the solution of the basic TLS problem 6.2.2.

Theorem 6.2.1 (Solution of the basic TLS problem $Ax \approx b$). *Let us take the singular value decomposition:* $[A; b] = U \Sigma V^T$ *where* $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n, \sigma_{n+1})$. *Let* $\hat{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n, 0$. *If* $\sigma_n > \sigma_{n+1}$ *and* $v_{n+1,n+1} \neq 0$ *then*

$$[\Delta \hat{A}; \Delta \hat{b}] = [A; b] - [\hat{A}; \hat{b}] = \sigma_{n+1} u_{n+1} v_{n+1}^T, \tag{6.18}$$

and

$$[\hat{A}; \hat{b}] = U \hat{\Sigma} V^T \tag{6.19}$$

solves the TLS problem. Moreover,

$$\hat{x} = -\frac{1}{v_{n+1,n+1}} [v_{1,n+1}, \ldots, v_{n,n+1}]^T \tag{6.20}$$

exists and is the unique solution to $\hat{A}x = \hat{b}$ *[Van Huffel and Vandewalle, 1991].*

As one can see solving the TLS problem is practically equivalent with computing a singular value decomposition. Unfortunately, for this version of the TLS we require that $v_{n+1,n+1} \neq 0$ and $\sigma_n > \sigma_{n+1}$. But, in the case of frequency domain system identification an alternative version presented in the followings is needed where we do not have this problem.

Problem 6.2.3 (Orthogonal l_2 approximation problem). *Consider the data matrix* $A \in \mathbb{R}^{m \times n}$. *The orthogonal approximation problem seeks to*

$$\underset{x \in \mathbb{R}^n}{\mathrm{minimize}} \|Ax\|_2 \text{ subject to } x^T x = 1 \tag{6.21}$$

[Van Huffel and Vandewalle, 1991].

The vector e defined as $e = Ax$ can be considered as a vector of residuals or errors on each equation. Therefore, the orthogonal l_2 approximation problem can also be interpreted as an equation error identification. Obviously, using the square of the roots leads to an equivalent problem:

$$\underset{x \in \mathbb{R}^n}{\mathrm{minimise}} \|Ax\|_2^2 \text{ subject to } x^T x = 1. \tag{6.22}$$

6.2.2.3   TLS in Frequency Domain System Identification

The orthogonal l_2 approximation problem is used in the frequency domain system identification. The basic approximation problem is

$$\frac{N(\Omega, \theta)}{D(\Omega, \theta)} \approx \frac{Y(\Omega)}{U(\Omega)}. \tag{6.23}$$

Rewriting the equation (6.23) results

$$N(\Omega, \theta)U(\Omega) - D(\Omega, \theta)Y(\Omega) \approx 0. \tag{6.24}$$

The left side is linear in the parameter vector θ because of the definitions (2.46) and (2.47). For every $k = 1, \ldots, F$ it gives an equation which can be collected in an over-determined linear equations of form (6.12). As in the theorem 6.2.3 the notation A is used, so the over-determined linear equations can be written as

$$A\theta = 0 \tag{6.25}$$

where the elements of A are

$$A_{k,l} = U(\Omega_k)\Omega_k^{l-1} \text{ if } l \leq n_\alpha + 1,$$

$$A_{k,l} = -Y(\Omega_k)\Omega_k^{l-1-(n_\alpha+1)} \text{ if } l > n_\alpha + 1. \tag{6.26}$$

Using (6.1) the noise of A can be introduced:

$$A = A_0 + N_A \tag{6.27}$$

where the noise matrix N_A can be generated like in (6.26):

$$N_{A,k,l} = N_U(\Omega_k)\Omega_k^{l-1} \text{ if } l \leq n_\alpha + 1,$$

$$N_{A,k,l} = -N_Y(\Omega_k)\Omega_k^{l-1-(n_\alpha+1)} \text{ if } l > n_\alpha + 1. \tag{6.28}$$

Problem 6.2.4 (TLS problem in frequency domain system identification)**.** *The TLS problem in frequency domain identification seeks to*

$$\underset{\hat{A}}{\text{minimize}} \|(A - \hat{A})\|_F^2 \tag{6.29}$$

subject to

$$\hat{A}\theta = 0 \text{ and } \theta^T\theta = 1. \tag{6.30}$$

Applying the modification (6.22) of the theorem 6.2.3 it can be proven that minimiza-

tion of the following cost function is equivalent to solving the problem 6.2.4 [Pintelon et al., 1994]:

$$C_{\text{TLS}} = \frac{\sum_{k=1}^{F} |N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)|^2}{\theta^T \theta}. \tag{6.31}$$

6.2.3 Generalized TLS

The original TLS problem can be generalized in different ways. A possible extension is to introduce a weighting function. In this thesis, this extension is not considered. An other possible generalization is to apply the covariance matrix in order that optimal noise weighting is obtained. This method is called generalized total least squares problem (GTLS) [Pintelon et al., 1998].

In GTLS problem the same error equations (6.24) is valid but its variance at every frequency Ω_k is used. At the frequency Ω_k the variance of the left side of (6.24) is

$$\mathbb{E}\left\{N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)^* N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)\right\}$$

$$= |N(\Omega_k, \theta)|^2 \mathbb{E}\{U^*(\Omega_k)U(\Omega_k)\} + |D(\Omega_k, \theta)|^2 \mathbb{E}\{Y^*(\Omega_k)U(\Omega_k)\}$$

$$= |N(\Omega_k, \theta)|^2 \sigma_{U,k}^2 + |D(\Omega_k, \theta)|^2 \sigma_{Y,k}^2. \tag{6.32}$$

The variance (6.32) establishes the variance of element of N_A. Hence, a covariance matrix C can be constructed:

$$C_{k,l} = \sigma_{U,k}^2 \Omega_k^{l-1} \text{ if } l \leq n_\alpha + 1$$

$$C_{k,l} = -\sigma_{Y,k}^2 \Omega_k^{l-1-(n_\alpha+1)} \text{ if } l > n_\alpha + 1. \tag{6.33}$$

Problem 6.2.5 (GTLS problem in frequency domain system identification). *The GTLS problem in frequency domain identification seeks to*

$$\underset{\hat{A}}{\text{minimize}} \|(A - \hat{A})C^{-1}\|_F^2 \tag{6.34}$$

subject to

$$\hat{A}\theta = 0 \text{ and } \theta^T \theta = 1. \tag{6.35}$$

The equivalent cost function of the GTLS problem:

$$C_{\text{GTLS}} = \frac{\sum_{k=1}^{F} |N(\Omega_k, \theta)U(\Omega_k) - D(\Omega_k, \theta)Y(\Omega_k)|^2}{\sum_{k=1}^{F} (\sigma_{U,k}^2 |N(\Omega_k, \theta)|^2 + \sigma_{Y,k}^2 |D(\Omega_k, \theta)|^2 - 2\text{Re}\{\rho_{YU,k}D(\Omega_k, \theta)N^*(\Omega_k, \theta)\})}. \tag{6.36}$$

It can be seen that this cost function is very similar to the maximum likelihood cost function (6.5).

6.3 Parameter Transformation

A transformation is nothing else than a function from the original parameter space to a new one. There are more possible reasons of a transformation: increase the numerical stability, reducing the number of the parameters, etc. In practise the transformation of the parameter vector is performed in many cases, although this is not always noticed. This chapter covers three cases:

- frequency scaling,

- orthogonal polynomials,

- known subsystem.

Frequency scaling and orthogonal polynomials are used only in the s-domain, although generally these transformations can be evaluated in the z-domain, too. The case of known subsystem can be applied in both domains.

The transformations can be combined and since the presented transformations are linear, it can be evaluated by multiplying matrices.

6.3.1 Frequency Scaling

In an s-domain identification problem the length of the frequency interval can be so wide that models even of moderate orders cannot be handled because of numerical imprecision [Pintelon et al., 1994].

A possible solution is to divide every frequency by a constant and therefore the frequency interval may be narrowed. The numerical stability of the TLS problem is dependent on the condition number of A. However, the optimal frequency scaling, when the condition number is minimal, is not independent of the model. Therefore, in practise using the arithmetic mean of the frequency set seems to be a good solution. The scaling factor is

$$\omega_{\text{scale}} = \frac{\omega_{\min} + \omega_{\max}}{2} \tag{6.37}$$

and $N(j\omega/\omega_{\text{scale}}, \theta)$ and $D(j\omega/\omega_{\text{scale}}, \theta)$ are used instead of $N(j\omega, \theta)$ and $D(j\omega, \theta)$. Essentially the bandpass spectrum is moved to the radian frequency 1. Thus, the parameter vector is scaled. In some cases, the median value of the applied frequency set is used as the scale factor [Pintelon and Kollár, 2004], [Pintelon et al., 2006]

$$\omega_{\text{scale}} = \text{median}\{\omega_1, \ldots, \omega_F\}. \tag{6.38}$$

In general the scaling factor can be any non zero number. Moreover, every power of Ω in the polynomials can be scaled independently by arbitrary numbers. For example, the logarithm of the the frequency $\log \omega^k$ can be applied.

The relationship between $N(j\omega/\omega_{\mathrm{scale}}, \theta)$ and $N(j\omega, \theta)$ is linear, and a transformation matrix of the parameter vector θ can be introduced:

$$N(j\omega/\omega_{\mathrm{scale}}, \theta) = N(j\omega, T_{N,\mathrm{scale}}\theta) \tag{6.39}$$

where

$$T_{N,\mathrm{scale}} = \mathrm{diag}\left(1, \frac{1}{\omega_{\mathrm{scale}}}, \ldots, \frac{1}{\omega_{\mathrm{scale}}^{n_\beta}}\right). \tag{6.40}$$

Similarly,

$$D(j\omega/\omega_{\mathrm{scale}}, \theta) = D(j\omega, T_{D,\mathrm{scale}}\theta) \tag{6.41}$$

where

$$T_{D,\mathrm{scale}} = \mathrm{diag}\left(1, \frac{1}{\omega_{\mathrm{scale}}}, \ldots, \frac{1}{\omega_{\mathrm{scale}}^{n_\beta}}\right). \tag{6.42}$$

Finally, the transformation matrix of the parameter vector can be composed as

$$T_{\mathrm{scale}} = \begin{bmatrix} T_{D,\mathrm{scale}} & 0 \\ 0 & T_{N,\mathrm{scale}} \end{bmatrix}. \tag{6.43}$$

6.3.2 Orthogonal Polynomials

The usual polynomial basis is $1, x, x^2, \ldots, x^r$. The numerical stability can be increased by the use of so-called orthogonal polynomials [Rolain et al., 1995a], [Rolain et al., 1995b]. Polynomials of the transfer function (2.46) can be expressed in a new basis

$$H(\Omega, \theta) = \frac{N(\Omega, \theta)}{D(\Omega, \theta)} = \frac{\sum_{r=0}^{n_b} b_r q_r(\Omega)}{\sum_{r=0}^{n_a} a_r p_r(\Omega)} \tag{6.44}$$

where $p_r(\Omega)$ and $q_r(\Omega)$ are polynomials of order r.

The orthogonal polynomials are used in the case of finite frequency grid and in the case of s-domain. Let us denote ω_k $(k = 1, \ldots, F)$ is the kth element of the frequency grid. F is cardinality of the frequency set. The Forsythe polynomial basis is generated using the following recursion formula:

$$R_r(x) = Z_r\left(xR_{r-1}(j\omega_k) + Z_{r-1}R_{r-2}(x)\right)$$

$$Z_r^2 = 2\sum_{k=1}^{F} |R_r(j\omega_k)|^2 |W(\omega_k)|^2 \tag{6.45}$$

where $W(\omega_k)$ is the discrete weighting function [Rolain et al., 1995a]. The first two poly-

nomials are defined as follows:

$$R_0(x) = \frac{1}{Z_1}, \quad R_1(x) = \frac{1}{Z_0 Z_1} x. \tag{6.46}$$

The orthonormal bases of the numerator and the denominator are computed separately because different sub-matrices in (6.24) belong to the numerator and the denominator. In the case of the numerator the weighting $W(\omega_k)$ equals with $U(j\omega)$ and in the case of denominator $W(\omega_k) = Y(j\omega_k)$. To compute the orthogonal polynomials in (6.45) $j\omega_k$, $k = 1, \ldots, F$ are substituted for x. It means that in the end the following two sub-matrices are formulated

$$B_n = [R_{n,0}, R_{n,1}, \ldots, R_{n,n_\beta}] \in \mathbb{C}^{F \times (n_\beta+1)} \tag{6.47}$$

and

$$B_d = [R_{d,0}, R_{d,1}, \ldots, R_{d,n_\beta}] \in \mathbb{C}^{F \times (n_\alpha+1)} \tag{6.48}$$

where $R_{n,k}$ and $R_{d,k}$ denote the kth orthogonal polynomial of the numerator and denominator, respectively. Note that $R_{2r}(x)$ and $R_{2r+1}(x)$ are respectively even and odd polynomials in x. The evaluation of these polynomials has to occur in a numerically stable way.

Since writing polynomials in a new basis is linear transformation, it is important to note that the transformation polynomials to orthogonal polynomials is a *linear* one. Hence, it is equivalent to a transformation of the parameter vector. In this case the transformation matrix is denoted by T_{orthpol}.

6.3.3   Known subsystem

Another possible application of the corrected TLS algorithm arrives when a part of the DUT is known *before* the estimation process. The measurement setup is in figure 6.2.

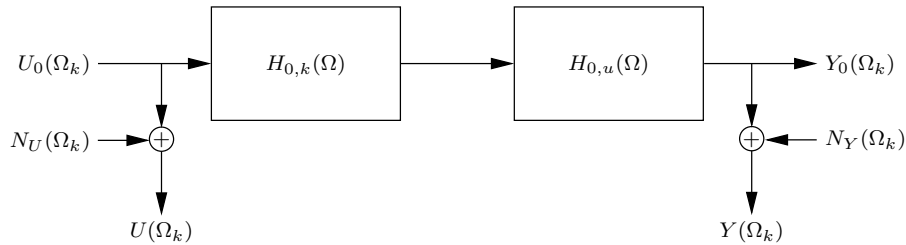Figure 6.2: Frequency domain representation of the measurement process in the case of known subsystem. $U_0(\Omega_k)$ and $Y_0(\Omega_k)$ are the real input and output, respectively. $U(\Omega_k)$, $Y(\Omega_k)$ are the measured quantities. $N_U(\Omega_k)$, $N_Y(\Omega_k)$ denote the noise processes. $H_{0,k}(\Omega)$ is the known subsystem and $H_{0,u}(\Omega)$ is the unknown part of the transfer function.

The transfer function $H_0(\Omega)$ can be written as

$$H_0(\Omega) = H_{0,u}(\Omega) H_{0,k}(\Omega) \qquad (6.49)$$

where $H_{0,u}(\Omega)$ is the unknown and $H_{0,k}(\Omega)$ is the known subsystem. Orders of the transfer function are not increased and disregarding the trivial cases the length of the parameter vector is decreased. In the new framework new notations is introduced for the numerator, the denominator and for the parameter vector: $\tilde{N}(\Omega, \theta)$, $\tilde{D}(\Omega, \theta)$, $\tilde{\theta}$.

The new form of (6.24) is

$$\tilde{N}(\Omega, \tilde{\theta}) U(\Omega) H_{0,k}(\Omega) - \tilde{D}(\Omega, \tilde{\theta}) Y(\Omega) \approx 0. \qquad (6.50)$$

or if the corresponding part of the parameter vector of $H_0(\Omega)$ is known:

$$\tilde{N}(\Omega, \tilde{\theta}) N_{0,k}(\Omega, \theta_k) U(\Omega) - \tilde{D}(\Omega, \tilde{\theta}) D_{0,k}(\Omega, \theta_k) Y(\Omega) \approx 0 \qquad (6.51)$$

where θ_k denotes the parameter vector of the known part, $N_{0,k}(\Omega, \theta_k)$ and $D_{0,k}(\Omega, \theta_k)$ are the corresponding numerator and denominator, respectively.

It can be seen that (6.50) and (6.51) determine a new matrix $\tilde{A}$ which has fewer columns than A has. Moreover, there is a linear relation between $\tilde{N}(\Omega, \tilde{\theta}) N_{0,k}(\Omega, \theta_k)$ and $N(\Omega, \theta)$. The linear map can be extended for $\tilde{D}(\Omega, \tilde{\theta}) D_{0,k}(\Omega, \theta_k)$ and $D(\Omega, \theta)$. The transformation $T_{\text{known subs.}}$ projects the parameter space onto a subspace:

$$T_{\text{known subs.}} \tilde{\theta} = \theta. \qquad (6.52)$$

6.4 Basis Transformation for the TLS Problem

As we have seen in the previous section sometimes transformation of the parameter vector is needed. Since every transformation used in this thesis is linear, it is very easy to combine them: transformation matrices must be multiplied. In the ultimate step the final transformation matrix T appears.

The reader can find one of the presented results of the thesis here. First, the cost function in which a problem appears is presented. The mentioned problem led to the recognition that transforming the parameter space without transforming the constraint the obtained parameter vector is not equivalent to the solution of the original TLS problem. Therefore, a new version of the TLS problem must be specified, and in the thesis the problem 6.4.1 is formulated. An equivalent problem 6.4.2 and the new cost function (6.59) are also presented. At the end, an example is given in order to demonstrate the theoretical result.

The TLS problem involves a constraint for the norm of the parameter vector. Simply transforming the parameter vector in (6.23) but not transforming the constraint leads to the following cost function

$$\frac{\sum_{k=1}^{F} |N(j\omega_k, T\theta) U(j\omega_k) - D(j\omega_k, T\theta) Y(j\omega_k)|^2}{\theta^T \theta}. \qquad (6.53)$$

Obviously solution of the minimization of (6.53) do not equal to the solution of the original TLS problem. Moreover, after minimization (6.53) the obtained parameter vector must be inverse-transformed which causes that the norm of ultimate parameter vector do not equal to 1.

Therefore a correction of the TLS problem is needed to overcome the mentioned problems. The effect of the transformation T in (6.25) results

$$AT\theta = 0. \tag{6.54}$$

The left side of (6.54) leads to the matrix A_m which is by definition $A_m = AT$.

Problem 6.4.1 (Basis transformation for the TLS problem in frequency domain system identification). *The corrected TLS problem in frequency domain identification seeks to*

$$\underset{\hat{A}_m}{\text{minimize}} \|(A_m - \hat{A}_m)\|_F^2 \tag{6.55}$$

subject to

$$\hat{A}_m\theta = 0 \text{ and } \theta^T T^T T\theta = 1 \tag{6.56}$$

where T is the corresponding transformation matrix.

There is an equivalent formulation:

Problem 6.4.2. *The corrected TLS problem in frequency domain identification seeks to*

$$\underset{\hat{A}_m}{\text{minimize}} \|(A_m - \hat{A}_m)T^{-1}\|_F^2 \tag{6.57}$$

subject to

$$\hat{A}_m\theta = 0 \text{ and } \theta^T \theta = 1. \tag{6.58}$$

The constraint is a bilinear expression. As a matter of fact (6.56) can be interpreted that the norm of θ equals to one, when the scalar product of the vector x_1 and the vector x_2 is defined as $x_1^T T^T T x_2$. Hence the corresponding cost function is

$$C_{\text{corrected TLS}} = \frac{\sum_{k=1}^{F} |N(j\omega_k, T\theta)U(j\omega_k) - D(j\omega_k, T\theta)Y(j\omega_k)|^2}{\theta^T T^T T\theta}. \tag{6.59}$$

It can be seen that even if T is not the identity matrix, the solution of the problem 6.4.1 is the same like the solution of the problem 6.2.4.

Figure 6.3 illustrates the different versions of the TLS problem. The solution of the corrected TLS problem can be computed by using the GSVD method.
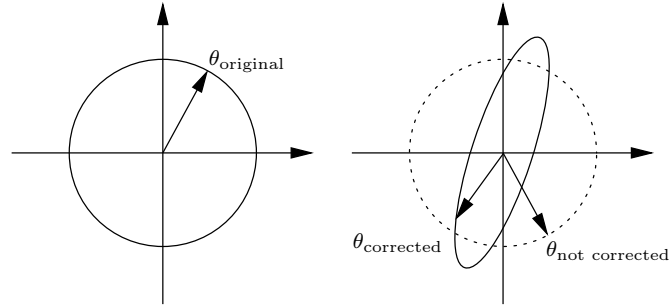
Figure 6.3: Illustration of the transformation of the constraint $\theta^T\theta = 1$. θ_{original} is the solution of the original TLS problem (left). $\theta_{\text{not corrected}}$ illustrates the solution when the constraint was not transformed and $\theta_{\text{corrected}}$ is equivalent with the original solution in the new basis.

6.4.0.1 GTLS

The GTLS problem presented in §6.2.3 does not need correction like the TLS problem. The reason is that in the GTLS problem the cost function (6.57) is replaced by

$$\underset{\hat{A}_m}{\text{minimize}} \|(A_m - \hat{A}_m)C_m^{-1}\|_F^2 \tag{6.60}$$

where C_m^{-1} is the inverse of the modified covariance matrix that can be computed (see (6.32))

$$C_m = TC. \tag{6.61}$$

The equation (6.61) shows that in the case of the GTLS problem the covariance matrix is also transformed, therefore it contains implicitly the transformation matrix T. And the modified formulation of the corrected TLS problem 6.4.1 shows that the constraint of the GTLS problem is implicitly transformed.

Another illustration of this effect is the denominators of the cost functions (6.31) and (6.36). In the denominator of the cost function of the GTLS problem contains the variance of the numerator, hence the transformed parameter vector.

6.4.1 Example

In the following tables some results obtained by running different algorithms are compared. The normalized difference vector will be used for this purpose. This means that if θ_1 and θ_2 denote the parameter vectors obtained as two estimation results, the expression of the normalized difference is

$$e = \frac{\| \theta_1 - \theta_2 \|_2}{\| \theta_1 \|_2}. \tag{6.62}$$

In our case $\| \theta_1 \|_2 > 0$.

Table 6.1: The normalized differences of the parameter vectors in the case of scaling.

scaling	compensation	e
yes	no	1.2384
yes	yes	0.0000

As an illustration of the procedure, the mechanical measurement of a robot arm is presented. The behavior of a flexible robot arm was measured by applying controlled torque to the vertical axis at one end of the arm, and measuring the tangential acceleration of the other end. The excitation signal was a multi-sine, generated with frequency components at $[1:2:199]df$, with $df = \frac{500}{4096} = 0.125$ Hz, that is, in the frequency range 0.125 Hz - 25 Hz. The originally flat multi-sine was distorted by the nonlinear behavior of the actuator. The odd harmonic frequencies provided that components produced by a squaring nonlinearity would not disturb the identification. The input and output signals were sampled with sampling frequency $f_s = 500$ Hz. Sampling was synchronized to the excitation signal so that 4096 samples were taken from each period. The data records contain 40960 points, that is, 10 periods were measured. Figure 6.4 shows the transfer function at the measured frequencies. The model is estimated with orders 4/6.
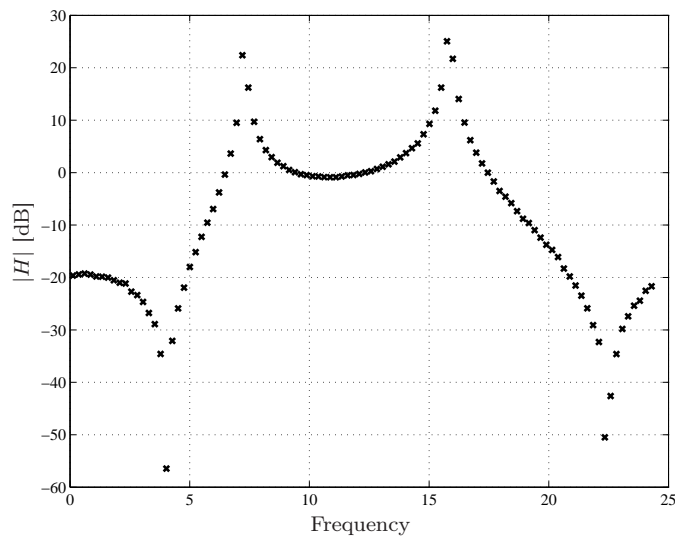
Figure 6.4: The measured transfer function of the robot arm.

The results of the frequency scaling are summarized in table 6.1. θ_1 is the parameter vector, which can be obtained without any transformation. In (6.62) θ_2 is the result of the estimation dappling a transformation. It can be the result of the TLS problem without correction or the estimation with correction.

The robot arm example is used again in order to illustrate the effects of the transformation to orthogonal polynomials. The normalized differences of the parameter vectors

Table 6.2: The normalized differences of the parameter vectors in the case of orthogonal polynomials.

representation	compensation	e
polynomials	no	1.2940
orthogonal polynomials	yes	0.0000

are listed in table 6.2.

We can see that the proposed correction of the TLS problem works well. The results coincide with the results without any transformation of the parameter vector.

6.5 Conclusion

In this chapter the correction of the TLS method is presented. The correction can be applied if the space of the parameter vector is transformed. It is shown how the norm constraint of the original problem needs to be transformed in order to obtain the same solution after the transformation.

The presented result fills a theoretical gap to complete the TLS theory. Although, the corrected and the non-corrected TLS estimators are inconsistent because the covariance matrices are not used, from a practical point of view, both methods can be used. It is also discussed that in the case of GTLS estimator the noise covariance matrix includes the transformation, so the constraint is implicitly transformed.

Chapter 7

Conclusions and Future Research Directions

There are several, still unsolved problems to be investigated in the future. Un-addressed and unsolved questions always remains when writing a thesis, but at some point the work should stop, because there is only a limited time for the research. Hereby I tried to summaries the open problems, which I could not deal with, because of time constraints.

Like the whole thesis, conclusions are divided into two parts. The first is about the stable approximation and the second one is about the TLS method reviewing conclusions and future research possibilities.

7.1 Stable Approximation with Additional Delay

7.1.1 Conclusions

The stable approximation problem is not trivial, even if the applied models are linear and time invariant. In this work L_2 approximations are studied. There are several published methods (some are presented in chapter 3) that are trying to solve non-linear optimization procedures. Practically, mathematics usually present only a simple existence theorem but there are only few theorems that guarantee the uniqueness or gives any useful information about the *global* optimum.

A possible solution of the stable approximation problem is if we allow to add some delay to the target function as described in §3.3.2. Of course, its application has limits, for example, it can be used only in open-loop configuration. An advantage of the method is that after adding a delay to the target function we seek the global minimum of the cost function what can be evaluated by effective algorithms. However, a disadvantage is that the appropriate delay value must be found.

In the thesis I could present a proof of this theorem for both s- and z-domain in the case of low order systems. The proof is valid if we apply a continuous frequency interval, i.e. the frequency domain is the imaginary axes or the unit circle. Hence, we can conclude that if the finite frequency grid is dense enough then the approximation problem can be very

close to the continuous one, therefore the minimum of the cost function with additional delay added to the target transfer function determines a stable system.

In addition to the theoretical results in chapter 5 we have shown and realized a new practical method based on numerical ODE solvers in order to find the appropriate delay. The proposed algorithm is tested with four already published examples. In one case the algorithm returned with the same result like the published one. In two cases the results are better than published designs, but in one case the method gives worse approximation than the one can be found in the literature.

7.1.2   Future Research Ideas

Although some open questions were answered, still there are open questions related to the studied problems. Some further questions that can be subject of further research are briefly the following:

- Can the proof be extended for higher order systems without the assumptions presented?

- Is there a counterexample with no stable approximation in the case of the finite frequency grid? The statement about the frequency grid which is becoming denser says nothing if we do not allow to extend the set of frequency points.

- In the proved theorem we used the fact that the approximation was over the whole imaginary axis or the whole unit circle. What is the situation in a band-limited application?

- Is there any method which is able to estimate the necessary orders of an approximation which is stable and the approximation error is less than a given level?

- In the thesis only the simplest ODE solvers are presented and realized. Have other methods (adaptive algorithms, methods for boundary problems) practical advantages?

- In the case of the differentiator the presented algorithm returns worse approximation than the one published in [Pintelon and Schoukens, 1990]. How should one improve the treated algorithm in order to avoid situations like this?

- Can the theoretical results be generalized for MIMO systems?

- Can the practical results be applied in the case of MIMO systems?

- The slightly non-linear systems can be defined by Volterra series. Is it possible to generalize the results to Volterra series?

7.2 Basis Transformation for the TLS Method

7.2.1 Conclusions

In this part of the thesis we addressed problems of the TLS method if the parameter vector is transformed.

The transformation maps a finite dimensional linear space to a finite dimensional linear space, therefore a matrix can be associated. In s-domain system identification without frequency scaling it is almost impossible to numerically handle the ill-conditioned matrices. And we showed that the frequency scaling and the usage of orthogonal polynomials transform the parameter vector.

The TLS problem in frequency domain system identification means that an over-determined linear equations has to be solved imposing a constraint for the length of the parameter vector. It turned out that solving the original TLS problem in the transformed base needs a transformation of the constraint. Chapter 6 presented the solution of this correction.

The case of GTLS estimator was also studied and we showed that the constraint of the GTLS problem implicitly contains the necessary transformation, so no correction is needed.

7.2.2 Future Research Ideas

Possible further question which it might be worth investigating is

- How can the presented correction be extended to the cases of multivariate input multivariate output systems?

Bibliography

Ansari, R. (1987). IIR discrete-time Hilbert transformers, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **35**: 1116–1119.

Baratchart, L., Leblond, J. and Partington, J. R. (1996). Hardy Approximation to L^∞ functions on Subset of the Circle, *Constructive Approximation* **12**: 423–436.

Baratchart, L., Leblond, J., Partington, J. R. and Torkhani, N. (1997). Robust Identification from Band-Limited Data, *IEEE Transaction on Automatic Control* **42**(9): 1318–1325.

Baratchart, L., Olivi, M. and Wielonsky, F. (1992). On a Rational Approximation Problem in the Real Hardy Spaces H_2, *Theoretical Computer Science* **12**: 175–197.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*, Cambridge University Press.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.*, McGraw-Hill, New York.

Cartwright, J. H. E. and Piro, O. (1992). The dynamics of Runge-Kutta methods, *International Journal of Bifurcation and Chaos* **2**(3): 427–449.

Chen, X. and Parks, T. W. (1987). Design of FIR filters in the complex domain, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **35**: 144–153.

Chui, N. L. C. and Maciejowski, J. M. (1996). Realization of Stable Models with Subspace Methods, *Automatica* **32**(11): 1587–1595.

Coddington, E. A. and Levinson, N. (1995). *Theory of Ordinary Differential Equations*, McGraw-Hill Book Company, Inc., New York.

Deczky, A. G. (1972). Synthesis of Recursive Digital Filters Using the Minimum p-Error Criterion, *IEEE Transaction on Audio and Electroacustics* **20**(4): 257–263.

D'haene, T., Pintelon, R. and Vandersteen, G. (2006). An Iterative Method to Stabilise a Transfer Function in the s- and z-Domains, *IEEE Transactions on Instrumentation and Measurement* **55**: 1192–1196.

Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.

Ellacott, S. and Williams, J. (1976a). Linear Chebyshev Approximation in the Complex Plane Using Lawson's Algorithm, *Mathematics of Computation* **30**(133): 35–44.

Ellacott, S. and Williams, J. (1976b). Rational Chebyshev Approximation in the Complex Plane, *SIAM Journal of Numerical Analysis* **13**(3): 310–323.

Feller, W. (1068). *An Introduction to Probability Theory*, Wiley, New York.

Golub, G. H. and Loan, C. F. V. (1996). *Matrix Computations*, The Johns Hopkins University Press.

Guillaume, G., Pintelon, R. and Schoukens, J. (1992). Nonparametric Frequency Response Functions Estimators Based on Non-linear Averaging Techniques, *IEEE Transactions on Instrumentation and Measurement* **41**(6): 739–746.

Hartman, P. (2002). *Ordinary Differential Equations*, Society for Industrial and Applied Mathematics, Philadelphia.

Jacobson, N. (1985). *Basic Algebra I*, W.H. Freeman and Company.

Johnson, A. T. (1978). Simultaneous Magnitude and Phase Equalization Using Digital Filters, *IEEE Transaction on Circuits and Systems* **25**(5): 319–321.

Katayama, T. (2005). *Subspace Methods for System Identification*, Springer-Verlag.

Kollár, I. (1993). On Frequency-Domain Identification of Linear Systems, *IEEE Transactions on Instrumentation and Measurement* **42**(1): 2–6.

Kollár, I. (1994). *Frequency Domain System Identification Toolbox for Matlab*, The Mathworks, Inc., Natick.

Kollár, I., Pintelon, R. and Schoukens, J. (1990). Optimal FIR and IIR Hilbert Transformer Design via LS and Minimax Fitting, *IEEE Transactions on Instrumentation and Measurement* **39**(6): 847–852.

Kollár, I., Pintelon, R., Rolain, Y. and Schoukens, J. (1991). Correspondence: Another Step Towards an Ideal Data Acquisition Channel, *IEEE Transactions on Instrumentation and Measurement* **40**(3): 659–660.

Kollár, I., Pintelon, R., Rolain, Y. and Schoukens, J. (1990). Equalization of Data Acquisition Channels Using Digital Filters, *Periodica Polytachnica Ser. Electrical Engineering* **34**(3): 167–178.

Kryszig, E. (1979). *Advanced Engineering Mathematics*, Wiley, New York.

Ljung, L. (1993). Some Results on Identifying Linear Systems Using Frequency Domain Data, *Proceedings of the 32nd IEEE Conference on Decision and Control*, San Antonio, Texas, pp. 567–569.

Ljung, L. (1999). *System Identification - Theory for the User*, Prentice-Hall, Upper Saddle River, N J.

Moses, R. L. and Liu, D. (1991). Determining the Closest Stable Polynomial to an Unstable One, *IEEE Transaction on Signal Processing* **39**(4): 901–906.

Oppenheim, A. V. and Shafer, R. W. (1989). *Discrete-Time Signal Processing*, Prentice-Hall Inc.

Oppenheim, A. V., Willsky, A. S. and Young, I. T. (1983). *Signals and Systems*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, USA.

Pintelon, R. and Kollár, I. (2004). On the Frequency Scaling in Continuous-Time Modeling, *Proceedings of the 21st IEEE Instrumentation and Measurement Technology Conference (IMTC'2004)*, Como, Italy, pp. 1586–1589.

Pintelon, R. and Schoukens, J. (1990). Real-Time Integration and Differentiation of Analog Signals by Means of Digital Filtering, *IEEE Transactions on Instrumentation and Measurement* **39**(6): 923–927.

Pintelon, R. and Schoukens, J. (2001). *System Identification: A Frequency Domain Approach*, IEEE Press, Piscataway (USA).

Pintelon, R., Guillaume, P., Rolain, Y., Schoukens, J. and Van Hamme, H. (1994). Parametric Identification of Transfer Function in the Frequency Domain–A Survey, *IEEE Transactions on Automatic Control* **39**(11): 2245–2260.

Pintelon, R., Guillaume, P., Vandersteen, G. and Rolain, Y. (1998). Analysis, Development and Applications of TLS Algorithms in Frequency-Domain System Identification, *SIAM Journal on Matrix Analysis and Applications* **19**(4): 983–1004.

Pintelon, R., Rolain, Y., Vanden Bossche, M. and Schoukens, J. (1990). Towards an Ideal Data Acquisition Channel, *IEEE Transactions on Instrumentation and Measurement* **39**(1): 116–120.

Pintelon, R., Schoukens, J. and Guillaume, P. (2006). Continuous-Time Noise Modeling From Sampled Data, *IEEE Transactions on Instrumentation and Measurement* **55**(6): 2253–2258.

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (2002). *Numerical Recipes in C*, 2 edn, Cambridge University Press.

Rabiner, L. R. and Gold, B. (1975). *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ.

Rolain, Y., Pintelon, R., Xu, K. and Vold, H. (1995a). Best Conditioned Parametric Identification of Transfer Function Models in the Frequency Domain, *IEEE Transaction on Automatic Control* **40**(11): 1954–1960.

Rolain, Y., Schoukens, J. and Pintelon, R. (1995b). Order Estimation for Linear Time-Invariant Systems Using Frequency Domain Identification Methods, *Proceedings of 34th IEEE Conference on Decision and Control*, New Orleans, USA, pp. 3588–3593.

Rudin, W. (1976). *Principles of Mathematical Analysis*, McGraw-Hill Education, Europe.

Rudin, W. (1987). *Real and Complex Analysis*, McGraw Hill.

Rudin, W. (1991). *Functional Analysis*, McGraw Hill.

Sahoo, P. and Riedel, T. (1998). *Mean Value Theorems and Functional Equations*, World Scientific.

Sain, M. and Massey, J. (1969). Invertibility of Linear Time-Invariant Dynamical Systems, *IEEE Transaction on Automatic Control* **14**(2): 141–149.

Schetzen, M. (1980). *The Volterra and Wiener Theories of Nonlinear Systems*, John Wiley and Sons, New York.

Schnell, L. (ed.) (1993). *Technology of Electrical Measurements*, John Wiley and Sons, Chichester.

Schoukens, J. and Pintelon, R. (1991). *Identification of Linear Systems: a practical guideline to accurate modeling*, Pergamon Press, London.

Schoukens, J., Pintelon, R. and Van Hamme, H. (1994). Identification of Linear Dynamic Systems Using Piecewise Constant Excitations: Use, Misuse and Alternatives, *Automatica* **30**(7): 1153–1169.

Silverman, L. (1969). Inversion of Multivariable Linear Systems, *IEEE Transaction on Automatic Control* **14**(3): 270–276.

Söderström, T. and Stoica, P. (1989). *System Identification*, Prentice Hall Internation, UK.

Tanaka, H. and Katayama, T. (2005). Stochastic subspace identification guaranteeing stability and minimum phase, *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic.

Titchmarsh, E. (1937). *Introduction to the Theory of Fourier Integrals*, Clarendon Press.

Van Gestel, T., Suykens, J. A. K., Van Dooren, P. and De Moor, B. (2001). Identification of Stable Models in Subspace Identification by Using Regularization, *IEEE Transaction on Automatic Control* **46**(9): 1416–1420.

Van Huffel, S. and Vandewalle, J. (1991). *The Total Least Squares Problem - Computational Aspects and Analysis*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Van Overschee, P. and De Moor, B. (1996). *Subspace Identifiaction for Linear Systems: Theory, Implementation, Applications*, Boston, MA: Kluwer.

Vuerinckx, R. (1998). *Design of Digital Chebyshev Filters in the Complex Domain*, PhD thesis, Vrije Universiteit Brussel.

Vuerinckx, R., Rolain, Y., Schoukens, J. and Pintelon, R. (1996). Design of Stable IIR Filters in the Complex Domain by Automatic Delay Selection, *IEEE Transaction on Signal Processing* **44**(9): 2339–2344.

Zhou, K., Doyle, J. C. and Glover, K. (1995). *Robust and Optimal Control*, Prentice-Hall Inc., New Jersey, USA.

Zwillinger, D. (1997). *Handbook of Differential Equations*, Academic Press, Bolton.

Appendix A

Original Contributions

The achieved results are collected into three statements.

Statement 1: *I have proved the following existence theorem: if $n_\beta + n_\alpha \leq 2$ then for every complex target function a delay value τ exists with the following property: when the target function is modified by a factor of $e^{-j\omega\tau}$, minimization of the cost function*

$$C(\theta, \tau) = \int_I |T(\Omega)e^{-j\omega\tau} - H(\Omega, \theta)|^2 d\Omega \qquad (A.1)$$

yields a stable model.

The precise statement can be found in §4.3, on page 41. Chapter 4 on pp. 33–127 proves the theorem.

Existence statements are not useful unless the value can be found with reasonable computer time. This is assured by the following statement.

Statement 2: *I have developed a new algorithm to find such a delay value. The proposed algorithm gives better values (in terms of the value of the cost function) than previous proposals published in literature.*

The algorithm is based on numerical methods for solving ordinary differential equations. The new algorithm is described and analyzed in §5.6, in pp. 140–143. Theoretical considerations are in §5.6.2 in pp. 143–144. The flowchart of the method is in §5.6.3, in pp. 144–145. The §5.7 in pp. 145–157 contains numerical examples.

Statement 3: *I have recognized and analyzed how the transformation of the parameter vector changes the solution of the TLS method in frequency domain system identification. I gave the correcting transformation and proved that it indeed lets the original problem be solved.*

I have demonstrated the use of the correction for transformations used in practice.

The new result is treated in §6.4, in pp. 169–171. The parameter transformations illustrating the method are listed in §6.3, in pp. 166–169.

Appendix B

List of Publications

Papers in Periodicals in English

[1] L. Balogh and I. Kollár, "Generalization of a Total Least Squares Problem in Frequency Domain System Identification", *IEEE Transaction on Instrumentation and Measurement*, vol. 51, pp. 1353-57, Dec. 2002.

[2] B. Eged and L. Balogh, "Analytical Calculation of the Impedance of Lossy Power/Ground Planes", *IEEE Transaction on Instrumentation and Measurement*, vol. 52. pp. 1886-1891, Dec. 2003.

[3] L. Balogh and R. Pintelon, "Stable approximations of unstable transfer function models", *IEEE Transaction on Instrumentation and Measurement*, vol. 57 pp. 2720-2726., Dec. 2008.

Papers in International Conference Proceedings

[4] L. Balogh and I. Kollár, "On the Setting of Initial Values for the Iterative Solution of the Frequency Domain System Identification Problem", in *Proceedings of the IEEE International Symposium on Intelligent Signal Processing (WISP'1999)*, pp. 126–130., Budapest, Hungary, 4–7 Sep., 1999.

[5] L. Balogh and I. Kollár, "Generalization of a Total Least Squares Problem in Frequency Domain System Identification", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTC'2001)*, Budapest, Hungary, 21–23 May, 2001.

[6] B. Eged and L. Balogh, "Analytical Calculation of the Impedance of Lossy Power/Ground Planes", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTC'2001)*, Budapest, Hungary, 21–23 May, 2001.

[7] B. Eged and L. Balogh, "Impedance Calculation of Lossy Power/Ground Planes with Concentrated Parameter Model", *Polish-Czech-Hungarian Workshop on Circuit The-*

ory, Signal Processing, and Telecommunication Networks, Budapest, Hungary, 14–17 Sep., 2001.

[8] L. Balogh, I. Kollár and G. Gueret, "Variance of Fourier Coefficients Calculated from Overlapped Signal Segments for System Identification", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTC'2002)*, Anchorage, AL, USA, 21–23 May, 2002.

[9] L. Balogh and I. Kollár, "Angle of Arrival Estimation Based on Interferometer Principle", in *Proceedings of the IEEE International Symposium on Intelligent Signal Processing (WISP'2003)*, Budapest, Hungary, 4–6 Sept. 2003, pp. 219-24. WISP/2003, Budapest, Sep. 2003

[10] L. Balogh, and B. Eged, "Precision Direction Finding Algorithm based on Interferometer Structure", *Proceedings of the 11th Microcoll*, Budapest Sep. 2003.

[11] B. Eged, L. Balogh and D. Tóth, "FFT-based Digital Receiver Architecture for Fast-scanning Application", in *Proceedings of the AOC Roma Symposium and Expo*, Rome, Italy, May 2004.

[12] L. Balogh and R. Pintelon, "Stable Approximation of Unstable Transfer Function Models", in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference (IMTC'2006)*, Sorrento, Italy, April 24-27, 2006.

[13] L. Balogh, B. Fodor, A. Sárhegyi, I. Kollár, "Maximum Likelihood Estimation of ADC Parameters from Sine Wave Test Data", in *Proceedings of 15th IMEKO TC4 Symposium on Novelties in Electrical Measurements and Instrumentation and 12th Workshop on ADC Modelling and Testing (IMEKO'2007)*, Iasi, Romania, Sep. 19-21, 2007.

Presentations at Hungarian Conferences

[14] L. Balogh, "Iteratív algoritmusok kezdeti érték beállítása (On the Setting of Initial Values of Iterative Algorithms)", in *Végzős konferencia (Conference for Graduated Students)*, Budapest, Hungary, 26 Apr. 2000, in Hungarian.

[15] L. Balogh, "Generalization of a Total Least Squares Problem in Frequency Domain System Identification", in *Proceedings of the 8th PhD Mini-Symposium.* Budapest, Hungary: Budapest University of Technology and Economics, Department of Measurement and Information Systems, pp. 50-51, 1 Jan.–1 Feb. 2001.

[16] L. Balogh, "Stable Approximation with System Delay", in *Proceedings of the 9th PhD Mini-Symposium.* Budapest, Hungary: Budapest University of Technology and Economics, Department of Measurement and Information Systems, pp. 50-51, 4–5 Feb. 2002.

[17] L. Balogh, "Stable Approximation with Additional Delay in Frequency Domain", in *Proceedings of the 10th PhD Mini-Symposium.* Budapest, Hungary: Budapest University of Technology and Economics, Department of Measurement and Information Systems, pp. 40-41, 4–5 Feb. 2003.

Lecture Note

[18] J.G. Németh., G. Péceli, I. Kollár, L. Sujbert and L. Balogh, "Digitális jelfeldolgozás (Digital Signal Processing, Lecture Note)", *Department of Measurement and Information Systems*, Budapest, Hungary, in Hungarian.

System identification is a powerful technique for constructing accurate models of complex systems from noisy input-output observations. It mainly consists of three basic steps that are interrelated:

(1) design of the experiment;

(2) choice of a parametric model (black box or physical laws);

(3) and the estimation of the model parameters from noisy measurements.

According to the intended goal of the identification experiment - physical interpretation, simulation, prediction, or control - some additional properties may be imposed on the identified model such as reciprocity, passivity, stability, ...

This thesis presents both theoretical (theorems) and practical (algorithms) contributions to the third step of an identification experiment: the estimation of guaranteed stable models from noisy data. A two step procedure is proposed: in the first step an unconstrained model is identified from the noisy measurements. Next, if unstable, the unstable model is in a second step approximated by a guaranteed stable model by adding an appropriate delay to the target function. The final result is a stable model with bias and noise uncertainty bounds that is useful in open loop simulation or prediction applications.